

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 02:14:16 / Search time 1264.21 Seconds
(without alignments)
2383.909 Million cell updates/sec

Title: US-09-509-779-3_COPY_141_264

Perfect score: 124

Sequence: 1 CGAATCGTGGCCATCTGCA.....ATTCCTCCACAATGCTGC 124

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbun:*
3: em_estin:*
4: em_estmu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vr1:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	282	12	BM694092
2	124	100.0	360	12	BM705593
3	124	100.0	406	12	BM700705
4	124	100.0	416	9	AA258876

5	124	100.0	430	14	W38711
6	124	100.0	438	9	AA448759
7	124	100.0	460	9	AL702182
8	124	100.0	463	14	CA948789
9	124	100.0	469	9	AA010517
10	124	100.0	495	9	AM879993
11	124	100.0	531	13	BQ082699
12	124	100.0	542	12	BM619883
13	124	100.0	595	10	BF033587
14	124	100.0	600	12	B1668630
15	124	100.0	620	12	BM767171
16	124	100.0	621	13	BQ186247
17	124	100.0	622	14	CB217926
18	124	100.0	629	13	BQ632594
19	124	100.0	641	13	BQ632301
20	124	100.0	651	10	BG714665
21	124	100.0	653	13	BQ665507
22	124	100.0	654	12	BM979431
23	124	100.0	678	13	BU601136
24	124	100.0	702	9	A1338342
25	124	100.0	706	12	BM704660
26	124	100.0	715	10	BG718227
27	124	100.0	716	12	BG761806
28	124	100.0	745	10	BG121625
29	124	100.0	748	10	BG576768
30	124	100.0	754	12	BG766992
31	124	100.0	767	12	B1601470
32	124	100.0	793	10	BG111145
33	124	100.0	797	12	B1832643
34	124	100.0	797	12	B1856307
35	124	100.0	808	10	BG708518
36	124	100.0	815	12	B1457840
37	124	100.0	822	12	B1668735
38	124	100.0	827	12	B1828930
39	124	100.0	840	10	BG037022
40	124	100.0	847	12	BG773490
41	124	100.0	853	13	BQ876591
42	124	100.0	855	10	BE747000
43	124	100.0	856	10	BG753323
44	124	100.0	856	12	B1601855
45	124	100.0	864	12	B1757881

ALIGNMENTS

RESULT 1
LOCUS BM694092
DEFINITION UI-E-CII-af-o-15-0-UI r1 UI-E-CII Homo sapiens cDNA clone
UI-E-CII-af-o-15-0-UI 5', mRNA sequence.
ACCESSION BM694092.1 GI:19007350
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 282)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548

COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers
 1..282
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-af0-o-15-0-UI"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-C11"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is ACCTA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI)."

BASE COUNT 66 a 66 c 92 g 57 t 1 others
 ORIGIN

Query Match 100.0%; Score 124; DB 12; Length 282;
 Best Local Similarity 100.0%; Pred. No. 2,8e-54;
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTCTTGAATGCAAGCTGA 60
 DB 137 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTCTTGAATGCAAGCTGA 196
 OY 61 AAACAACAAGAGAGAGCTGTGTGGTCTGGGGAATGATATCTTCCCAACTG 120
 DB 197 AAACAACAAGAGAGAGCTGTGTGGTCTGGGGAATGATATCTTCCCAACTG 256
 OY 121 CTGC 124
 DB 257 CTGC 260

RESULT 2 360 bp mRNA linear EST 28-FEB-2002
 BM705593
 LOCUS
 DEFINITION UI-E-DW0-agi-c-06-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
 ACCESION BM705593
 VERSION BM705593.1 GI:19018851
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548

COMMENT Coordinated Laboratory for Computational Genomics

University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers
 1..360
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-DW0-agi-c-06-0-UI"
 /tissue_type="lens"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-DW0"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-DW0 is a cDNA library containing the following
 tissue(s): lens. The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was ligated to an EcoR I adaptor, digested with Not
 I, and cloned directionally into pT73-Pac vector. The
 oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is CGATTACCGA. This library
 was created for the program, Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI)."

BASE COUNT 90 a 79 c 106 g 85 t
 ORIGIN

Query Match 100.0%; Score 124; DB 12; Length 360;
 Best Local Similarity 100.0%; Pred. No. 3e-54;
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTCTTGAATGCAAGCTGA 60
 DB 54 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTCTTGAATGCAAGCTGA 113
 OY 61 AAACAACAAGAGAGAGCTGTGTGGTCTGGGGAATGATATCTTCCCAACTG 120
 DB 114 AAACAACAAGAGAGAGCTGTGTGGTCTGGGGAATGATATCTTCCCAACTG 173
 OY 121 CTGC 124
 DB 174 CTGC 177

RESULT 3 406 bp mRNA linear EST 28-FEB-2002
 BM700705
 LOCUS
 DEFINITION UI-E-CR1-aei-c-03-0-UI.r1 UI-E-CR1 Homo sapiens cDNA clone
 ACCESION BM700705
 VERSION BM700705.1 GI:19013963
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 JOURNAL discovery

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES
source

1. 406
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CRI-acl-c-03-0-UI"
/tissue_type="eye anterior segment"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CRI"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CRI is a normalized cDNA library containing the following tissue(s): eye anterior segment. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AATGCCGAT. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT
ORIGIN

90 a 104 c 133 g 79 t

Query Match 100.0%; Score 124; DB 12; Length 406;
Best Local Similarity 100.0%; Pred. No. 3.1e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTTCTTAGATGTCACACTGA 60
|||||
Db 178 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTTCTTAGATGTCACACTGA 237

QY 61 AAACAACAAGAGACTGTGTGTGTGTGGGAGATGATATCATTCCTCCACAACCTG 120
|||||
Db 238 AAACAACAAGAGACTGTGTGTGTGTGGGAGATGATATCATTCCTCCACAACCTG 297

QY 121 CTGC 124
|||||
Db 298 CTGC 301

RESULT 4
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AA258876 416 bp mRNA linear EST 13-AUG-1997
AA258876
263a01.r1 NCI CGAP GCB1 Homo sapiens cDNA clone IMAGE:66856 5'
similar to WP:ZK287.5 CE06614 ;, mRNA sequence.
AA258876.1 GI:1894001
EST.
Homo sapiens (human)
Homo sapiens

JOURNAL
MEDLINE
PUBMED
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 416)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapb-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 795 Std Error: 0.00
Seq primer: -28mJ rev2 ET from Amersham.

FEATURES
source

1. 416
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:66856"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI CGAP GCB1"
/note="Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Altman (NCI) and Dr. Gerald Marcu (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCACTGTAAGGAGGAGCGCGCCCTCTTTTCTTTT-3'
(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
ORIGIN

102 a 97 c 125 g 92 t

Query Match 100.0%; Score 124; DB 9; Length 416;
Best Local Similarity 100.0%; Pred. No. 3.1e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTTCTTAGATGTCACACTGA 60
|||||
Db 102 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTTCTTAGATGTCACACTGA 161

QY 61 AAACAACAAGAGACTGTGTGTGTGTGGGAGATGATATCATTCCTCCACAACCTG 120
|||||
Db 162 AAACAACAAGAGACTGTGTGTGTGTGGGAGATGATATCATTCCTCCACAACCTG 221

QY 121 CTGC 124
|||||
Db 222 CTGC 225

RESULT 5
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

W38711 430 bp mRNA linear EST 15-MAY-1996
W38711
2b27c08.r1 Soares parathyroid tumor NBHPA Homo sapiens cDNA clone
IMAGE:304814 5' similar to WP:F35012.9 CE00978 ;, mRNA sequence.
W38711.1 GI:1320487
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 430)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Matra, M., Parsons, J., Riekin, L., Rohlfing, T., Soares, M., Tan, F., Trevasaki, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.
The WashU-Merck Est Project

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by GfP (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
No 81 sequence available.
This clone (DKFZp68606155) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..460
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp68606155"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="686 (synonym: hicc3)"
/note="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiIb;
cDNA-collection"

BASE COUNT 105 a 114 c 148 g 93 t
ORIGIN

Query Match 100.0%; Score 124; DB 9; Length 460;
Best Local Similarity 100.0%; Pred. No. 3.2e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 60
166 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 225
DB 61 AAACAAACAGAGAGACTGTGTGTGCTGGGGAGATGAATCATCTCTCCACAACTG 120
226 AAACAAACAGAGAGACTGTGTGTGCTGGGGAGATGAATCATCTCTCCACAACTG 285
QY 121 CTGC 124
111
DB 286 CTGC 289

RESULT 8
LOCUS CA948789 463 bp mRNA linear EST 31-DEC-2002
DEFINITION iQ27905.Y1 HR85 islet Homo sapiens cDNA clone IMAGE: 5' similar to
TR:Q9Y5M7 Q9Y5M7 RING FINGER PROTEIN. [1]; mRNA sequence.
ACCESSION CA948789
VERSION CA948789.1 GI:27441666
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 463)
AUTHORS Melton,D., Brown,J., Kenty,G., Permut,A., Lee,C., Kaestner,K.,
Lemshke,I., Scaer,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Maria,M., Pepe,D., Wyllie,T., Martin,J., Blistein,A.,
Schmitt,A., Meising,B., Ritzer,E., Ronko,I., Bennett,J., Cardenas
, M., Gibbons,W., McCann,R., Cole,R., Tsagarashvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished
Other ESTs: iQ27905.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biochem.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: 40UP from Gibco
High quality sequence stop: 394.
Location/Qualifiers

FEATURES

source

1..463
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_1ib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permut Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

BASE COUNT 114 a 98 c 132 g 119 t
ORIGIN

Query Match 100.0%; Score 124; DB 14; Length 463;
Best Local Similarity 100.0%; Pred. No. 3.2e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 60
66 CGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 125
DB 61 AAACAAACAGAGAGACTGTGTGTGCTGGGGAGATGAATCATCTCTCCACAACTG 120
126 AAACAAACAGAGAGACTGTGTGTGCTGGGGAGATGAATCATCTCTCCACAACTG 185
QY 121 CTGC 124
111
DB 186 CTGC 189

RESULT 9
LOCUS AA010517 469 bp mRNA linear EST 09-MAY-1997
DEFINITION z109906.r1 Soares fetal_liver_spleen_infls_S1 Homo sapiens cDNA
clone IMAGE:430330.5' similar to WP:F35612.9 CH00978; mRNA
sequence.
ACCESSION AA010517
VERSION AA010517.1 GI:1471543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 469)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chisoe,S., Dietrich,N., Dubuque,T., Favellio,A., Gish,W., Hawkins
, M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
, B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierly-Meg,J., Trevaetis,E.,
Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
8889549
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the

source

1. .531
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S14K402-56-D03"
/cell_line="K402"
/lab_host="Top10F"
/clone_1ib="S14K402"
/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 120 a 141 c 161 g 109 t
ORIGIN

Query Match 100.0%; Score 124; DB 13; Length 531;
Best Local Similarity 100.0%; Pred. No. 3.4e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATCGTCCGCGCATCTGCAGGCTCCAGGTGATGATGCTCTTAAATGTCAGACTGA 60
DB 205 CGATACGTCGCCCATCTGCAGGCTCCAGGTGATGATGCTCTTAAATGTCAGACTGA 264
QY 61 AAACAACAAGAGAGACTGTGTGTGCTCGGGAGAAATGTAATCAATCTTCCACAACCTG 120
DB 265 AAACAACAAGAGAGACTGTGTGTGCTCGGGAGAAATGTAATCAATCTTCCACAACCTG 324
QY 121 CTGC 124
DB 325 CTGC 328

RESULT 12
BM761983 542 bp mRNA linear EST 04-MAR-2002
LOCUS K-EST0043062 S13KMS5 Homo sapiens cDNA clone S13KMS5-5-A10 5', mRNA
DEFINITION sequence.
ACCESSION BM761983
VERSION BM761983.1 GI:19091598
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 542)
AUTHORS Oh,K.U., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 5 row: A column: 10
High quality sequence stop: 542.
Location/Qualifiers
1. .542

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S13KMS5-5-A10"
/tissue_type="myeloma"
/cell_line="KMS-5"
/lab_host="Top10F"
/clone_1ib="S13KMS5"
/note="Vector: pCMS; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 125 a 137 c 164 g 116 t
ORIGIN

Query Match 100.0%; Score 124; DB 12; Length 542;
Best Local Similarity 100.0%; Pred. No. 3.4e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATCGTCCGCGCATCTGCAGGCTCCAGGTGATGATGCTCTTAAATGTCAGACTGA 60
DB 182 CGATCGTCCGCGCATCTGCAGGCTCCAGGTGATGATGCTCTTAAATGTCAGACTGA 241
QY 61 AAACAACAAGAGAGACTGTGTGTGCTCGGGAGAAATGTAATCAATCTTCCACAACCTG 120
DB 242 AAACAACAAGAGAGACTGTGTGTGCTCGGGAGAAATGTAATCAATCTTCCACAACCTG 301
QY 121 CTGC 124
DB 302 CTGC 305

RESULT 13
BF033587 595 bp mRNA linear EST 20-OCT-2000
LOCUS BF033587
DEFINITION 601453564P1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3857511 5', mRNA sequence.
ACCESSION BF033587
VERSION BF033587.1 GI:10741299
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 595)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cga@bbs-rt@mail.nih.gov
Tissue Procurement: DCTD/DBP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9587 row: 0 column: 16
High quality sequence stop: 595.
Location/Qualifiers
1. .595

FEATURES
source

			/organism="Homo sapiens"	
			/mol_type="mRNA"	
			/db_xref="taxon:9606"	
			/clone="IMAGE:3857511"	
			/tissue_type="adenocarcinoma"	
			/lab_host="DH10B (phage-resistant)"	
			/clone_lib="NIH MGC 66"	
			/note="Saliv; Vector: PCMV-SPORE6; Site_1_Next;	
			site_2: Oari; cloned unidirectionally. Primer: Oligo dT.	
			Average insert size 1.8 kb. Library constructed by Life	
			technologies."	
BASE COUNT	153	a	126	c 165 g 151 t
ORIGIN				
Query Match		100.0%;	Score 124;	DB 10; Length 595;
Best Local Similarity		100.0%;	Pred.No. 3,5e-54;	
Matches	124;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0
QY	1	CGATACGTGCGCCATCTGCAGGGATCCAGGTGATGATGTCGCTGCTTAGATGTTCAAGCTTA	60	
Db	96	CGATTGCGTGCGCATCTGTCAGGGATCCAGGTGATGATGTCGCTGCTTAGATGTTCAAGCTTA	155	
QY	61	AAACAACAAGAAGACTGTGTTGTGATCTGGAGGAAGATTAATCAATCCTTCACAACTG	120	
Db	156	AAACAACAAGAAGACTGTGTTGTGATCTGGAGGAAGATTAATCAATCCTTCACAACTG	215	
QY	121	CTGC	124	
Db	216	CTGC	219	

FEATURES	LOCATION	QUALIFIERS
RESULT 14	600 bp	mRNA
LOCUS	6032938343P1 NIH_MGC_96 Homo sapiens	linear EST 12-SEP-2001
DEFINITION	6032938343P1 NIH_MGC_96 Homo sapiens	cdna clone IMAGE:5313034 5',
ACCESSION	BI668630	mRNA sequence.
VERSION	BI668630.1	GI:15582863
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Enkayrta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.	
TITLE	1 (bases 1 to 600)	
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/ .	
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)	
	Unpublished	
	Contact: Robert Strausberg, Ph.D.	
	Email: cgabbs-remail.nih.gov	
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.	
	CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki	
	Toshiyuki and Piero Carninci (RIKEN)	
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)	
	DNA Sequencing by: Incyte Genomics, Inc.	
	Clone distribution: MGC clone distribution information can be	
	found through the I.M.A.G.E. Consortium/LNLN at:	
	http://image.lnl.gov	
	Plate: LLM11793	row: 6 column: 11
	High quality sequence stop: 596.	
	Location/Qualifiers	

```

FEATURES
source
    location/Qualifiers
    1. 600
        /organism="Homo sapiens"
        /mol_type="mrna"
        /db_xref="taxon:9606"
        /clone="IMAGE:5313034"
        /tissue_type="hypothalamus"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_96"
        /note="Organ: brain; Vector: pBluescript (modified
        pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
        ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
        size-selected for average insert size 2.3 kb and
        normalized to RQF 5. This is a primary library enriched

```

for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

Query Match	100.0%	Score 124	DB 12	Length 600
Best Local Similarity	100.0%	Pred. No. 3.5e-54		
Matches 124	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	CGATACGTGGCCACTTGTGCAGGGTCCAGGTCATGAGTCCCTGTCTTAGATGCAAGCTGA	60	
Db	169	CGATACGTGGCCACTTGTGCAGGGTCCAGGTCATGAGTCCCTGTCTTAGATGCAAGCTGA	228	
QY	61	AAACAAACAGAGGACTGTGTGTGTGTGGGGAGAAATGTAATCATTTCTTCCAAACTG	120	
Db	229	AAACAAACAGAGGACTGTGTGTGTGTGTGGGGAGAAATGTAATCATTTCTTCCAAACTG	288	
QY	121	CTGC 124		
Db	289	CTGC 292		

RESULT 15	
BM767171	
LOCUS	
DEFINITION	
BM767171	620 bp
KM-ESTD046611 S14K402 Homo sapiens	mRNA linear EST 04-MAR-2002
sequence.	cDNA clone S14K402-4-C07 5', mRNA

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 620)
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R. Oh,K.U., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE	21C Frontier Korean EST Project 2001
JOURNAL	Unpublished
COMMENT	Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel.: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 4 row: C column: 07
High quality sequence stop: 620.
location/Qualifiers
1. .620
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S14K402-4-C07"
/cell_line="K402"
/lab_host="Ttop10F"
/clone_1fb="S14K402"
/more="Organ: Stomach; Vector: pTZ19Rp1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then deapped with tabacco acid pyrophosphatase (TAP) . The decapped intact mRNA was ligated with DNA-RNA linker including EcoorI I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector" the dT-tailed vector was adjusted to have about 60nt . The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoorI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Bey method. The obtained cDNA vectors were used for transformation of

```

competent cells *E. coli* Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

BASE COUNT 145 a 154 c 177 g 144 t

Query Match 100.0%; Score 124; DB 12; Length 620;
Best Local Similarity 100.0%; Pred. No. 3.5e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTCGCGCATCTGCAGGGTCCAGTGATGATGCTGCTTGAATGTCACCTGA 60
DB 205 CGATACGTCGCGCATCTGCAGGGTCCAGTGATGATGCTGCTTGAATGTCACCTGA 264

QY 61 AAACAACAAGAGAGACTGTGTGTGTGGGGAGATGATTAATCATCTCTCCACAACTG 120
DB 265 AAACAACAAGAGAGACTGTGTGTGTGGGGAGATGATTAATCATCTCTCCACAACTG 324

QY 121 CTGC 124
DB 325 CTGC 328

RESULT 16 B0186247 621 bp mRNA linear EST 30-APR-2002
LOCUS B0186247
DEFINITION UI-E-EJ1-aj1-03-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone

ACCESSION B0186247
VERSION B0186247.1 GI:20361798
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 621)
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548

COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 REVERSE

FEATURES
source location/Qualifiers
1..621
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ1-aj1-03-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ1"
/note="Organ: eye, Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-EJ1 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes, AGAATCAAGA
; lens, CGATTAGCGA; eye anterior segment, AATGCCCAT;
optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and
Macular, GTCC; RPE and Choroid, ACCCTA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."

BASE COUNT 168 a 121 c 160 g 170 t 2 others

Query Match 100.0%; Score 124; DB 13; Length 621;
Best Local Similarity 100.0%; Pred. No. 3.5e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTCGCGCATCTGCAGGGTCCAGTGATGATGCTGCTTGAATGTCACCTGA 60
DB 20 CGATACGTCGCGCATCTGCAGGGTCCAGTGATGATGCTGCTTGAATGTCACCTGA 79

QY 61 AAACAACAAGAGAGACTGTGTGTGTGGGGAGATGATTAATCATCTCTCCACAACTG 120
DB 80 AAACAACAAGAGAGACTGTGTGTGTGGGGAGATGATTAATCATCTCTCCACAACTG 139

QY 121 CTGC 124
DB 140 CTGC 143

RESULT 17 CB217926 622 bp mRNA linear EST 06-FEB-2003
LOCUS CB217926
DEFINITION NISC nb05g10.y1 COGENE 6E MAN Homo sapiens cDNA clone IMAGE:5795514
5', mRNA sequence.

ACCESSION CB217926
VERSION CB217926.1 GI:28266118
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 622)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
cDNA Library Preparation:
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLNL12898 row: M column: 19
Seq primer: M13P1 reverse primer (ABI).

FEATURES
source location/Qualifiers
1..622
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5795514"
/tissue_type="mandible, pooled"
/dev_stage="embryo, 6 weeks postconception"
/lab_host="DH10B"
/clone_lib="COGENE 6E MAN"
/note="Vector: pAMP1; cDNA primed using oligo-dT primer,
directionally cloned into UDG sites of pAMP1. Size
selected for insert sizes ranging from 0.2-2.0 kb.

Normalized to Cof5. Primary library, non-amplified.
Library constructed by M. Lovett. For more information on
this library, please contact R. Tidwell (Washington
University) or visit the COGENE website at
<http://hg.wustl.edu/COGENE/>

BASE COUNT 144 a 147 c 182 g 149 t
ORIGIN

Query Match 100.0%; Score 124; DB 14; Length 622;
Best Local Similarity 100.0%; Pred. No. 3.5e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGATACGTCGCCCATCTGCAGAGGTCCAGTGTGATGCTCTTAAATGTCAGCTGA 60
|||
DB 165 CGATACGTCGCCCATCTGCAGAGGTCCAGTGTGATGCTCTTAAATGTCAGCTGA 224
|||
OY 61 AAACAACAGAGAGACTGTGTGTGTGTGGGAGATGTAATCATTCCTTCCACACTG 120
|||
DB 225 AAACAACAGAGAGACTGTGTGTGTGTGGGAGATGTAATCATTCCTTCCACACTG 284
|||
OY 121 CTGC 124
|||
DB 285 CTGC 288

RESULT 18 629 bp mRNA linear EST 02-JUL-2002
B0632594 1125c04.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6031086 5'

DEFINITION similar to TR:Q9Y5M7 Q9Y5M7 RING FINGER PROTEIN. [1] ;, mRNA
sequence.

ACCESSION B0632594 GI:21684112
VERSION B0632594.1
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 629)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Bistrain, A.,
Schmitt, A., Theising, B., Rittler, E., Ronko, I., Bennett, J., Cardenas
, M., Gibbons, M., McCann, R., Cole, R., Tsagarisvilli, R., Williams, T.,
Jackson, Y. and Bowers, Y.

AUTHORS

TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished
COMMENT Other ESTs: 1125c04.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@iobp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Glibco
High quality sequence stop: 495.
Location/Qualifiers

FEATURES
SOURCE 1..629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6031086"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_1ib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.

Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@im.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692.

BASE COUNT 157 a 141 c 173 g 158 t
ORIGIN

Query Match 100.0%; Score 124; DB 13; Length 629;
Best Local Similarity 100.0%; Pred. No. 3.5e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGATACGTCGCCCATCTGCAGAGGTCCAGTGTGATGCTCTTAAATGTCAGCTGA 60
|||
DB 126 CGATACGTCGCCCATCTGCAGAGGTCCAGTGTGATGCTCTTAAATGTCAGCTGA 185
|||
OY 61 AAACAACAGAGAGACTGTGTGTGTGTGGGAGATGTAATCATTCCTTCCACACTG 120
|||
DB 186 AAACAACAGAGAGACTGTGTGTGTGTGGGAGATGTAATCATTCCTTCCACACTG 245
|||
OY 121 CTGC 124
|||
DB 246 CTGC 249

RESULT 19 641 bp mRNA linear EST 02-JUL-2002
B0632301 1125c04.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6031086 3'

DEFINITION similar to TR:Q9Y5M7 Q9Y5M7 RING FINGER PROTEIN. [1] ;, mRNA
sequence.

ACCESSION B0632301 GI:21683819
VERSION B0632301
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 641)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Bistrain, A.,
Schmitt, A., Theising, B., Rittler, E., Ronko, I., Bennett, J., Cardenas
, M., Gibbons, M., McCann, R., Cole, R., Tsagarisvilli, R., Williams, T.,
Jackson, Y. and Bowers, Y.

AUTHORS

TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished
COMMENT Other ESTs: 1125c04.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@iobp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40UP from Glibco
High quality sequence stop: 449.
Location/Qualifiers

FEATURES
SOURCE 1..641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6031086"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_1ib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.

NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permut Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

BASE COUNT 176 a 159 c 125 g 181 t

ORIGIN

Query Match 100.0%; Score 124; DB 13; Length 641;
Best Local Similarity 100.0%; Pred. No. 3.5e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATAGTGGCCATCTGCAGGGTCCAGTGAATGCTGCTTAGATGTAACCTGA 60
|||||
Db 624 CGATAGTGGCCATCTGCAGGGTCCAGTGAATGCTGCTTAGATGTAACCTGA 565
|||||
QY 61 AACCAACAAGAGAGACTGTGTGTGTGGGAGATGTAATCATTCCTCCACAAC 120
|||||
Db 564 AACCAACAAGAGAGACTGTGTGTGTGGGAGATGTAATCATTCCTCCACAAC 505
|||||
QY 121 CTGC 124
|||||
Db 504 CTGC 501

RESULT 20
Bg714665 651 bp mRNA linear EST 08-MAY-2001
LOCUS 602677016F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799834 5',
DEFINITION mRNA sequence.

ACCESSION Bg714665
VERSION Bg714665.1 GI:139393596
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 651)
NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgaops-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHRRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10689 row: 0 column: 03
High quality sequence stop: 643.

FEATURES

source
1. 651
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4799834"
/issue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/note="Organ: Brain; Vector: pBluescript (modified
bluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHRRI, National

Institutes of Health). Note: this is a NIH_MGC library."

BASE COUNT 157 a 148 c 190 g 156 t

ORIGIN

Query Match 100.0%; Score 124; DB 10; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.5e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATAGTGGCCATCTGCAGGGTCCAGTGAATGCTGCTTAGATGTAACCTGA 60
|||||
Db 169 CGATAGTGGCCATCTGCAGGGTCCAGTGAATGCTGCTTAGATGTAACCTGA 228
|||||
QY 61 AACCAACAAGAGAGACTGTGTGTGTGGGAGATGTAATCATTCCTCCACAAC 120
|||||
Db 229 AACCAACAAGAGAGACTGTGTGTGTGGGAGATGTAATCATTCCTCCACAAC 288
|||||
QY 121 CTGC 124
|||||
Db 289 CTGC 292

RESULT 21
BU665507 653 bp mRNA linear EST 30-SEP-2002
LOCUS c1132b08.z1 Hembase; Erythroid Precursor Cells (LCB:c1 library)
DEFINITION Homo sapiens cDNA clone c1132b08 5', mRNA sequence.

ACCESSION BU665507
VERSION BU665507.1 GI:23377694
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 653)
Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished

Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jml7@nih.gov
The 'c1' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plate: 132 row: 0 column: 08
Seq primer: 5' lambda-TripLex2 Sequencing Primer.
Location/Qualifiers

FEATURES

source
1. 653
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c1132b08"
/sex="unknown"
/issue_type="Blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:c1
library)"
/note="Organ: blood; Vector: pTriplex2; Site 1: SfiI;
Site 2: SfiI; A complementary DNA (cDNA) library from
human erythroid precursor cells was constructed using
SMART PCR (polymerase chain reaction) cDNA library
construction kit (Clontech, Palo Alto, CA) according to
the manufacturer's directions, but with slight

modifications. Briefly, reverse transcription was performed in the presence of 1 umol/l peptide nucleic acid (PNA) oligos (N-terminal)-biotin-GTC-CAC-CCG-AGG-CCT-G-(C-terminal) and (N-terminal) -biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC; <http://www.nisc.nih.gov/>).

BASE COUNT 147 a 163 c 187 g 156 t
ORIGIN

Query Match 100.0%; Score 124; DB 13; Length 653;
Best Local Similarity 100.0%; Pred. No. 3.6e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTGGCCATCTGCAGGCTCCAGTGGATGATGCTGCTTGAATGTAAGCTGA 60
DB 220 CGATACGTGGCCATCTGCAGGCTCCAGTGGATGATGCTGCTTGAATGTAAGCTGA 279
QY 61 AAACAAACAGAGAGACTGTGTGTGTGTGGGAGATGTAATCATTCCTCCACAAC 120
DB 280 AAACAAACAGAGAGACTGTGTGTGTGTGGGAGATGTAATCATTCCTCCACAAC 339
QY 121 CTGC 124
DB 340 CTGC 343

RESULT 22
BM979431 654 bp mRNA linear EST 21-FEB-2003
LOCUS UI-CF-DUI-adr-f-23-0-UI-s1 UI-CF-DUI Homo sapiens cDNA clone
DEFINITION UI-CF-DUI-adr-f-23-0-UI-s1, mRNA sequence.
VERSION BM979431
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 654)
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548

COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.openbiosystems.com) or from Open Biosystems
(www.openbiosystems.com).
Seg primer: M13 FORWARD
POLYA=Yes

FEATURES
source Location/Qualifiers
1..654

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-adr-f-23-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"

/lab host="DH10B (Life Technologies) (T1 phage resistant)"
/clone.lib="UI-CF-DUI"
/note="Organ: Lung; Vector: pRT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-DUI is a normalized cDNA library containing the following tissue(s): Primary Lung Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pRT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (drr)18 tail. The sequence tag for this library is GGCTGTAGGC.
TAG_LIB=UI-CF-DUI
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_SEQ=GGCTGTAGGC"

BASE COUNT 173 a 156 c 127 g 198 t

Query Match 100.0%; Score 124; DB 12; Length 654;
Best Local Similarity 100.0%; Pred. No. 3.6e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTGGCCATCTGCAGGCTCCAGTGGATGATGCTGCTTGAATGTAAGCTGA 60
DB 643 CGATACGTGGCCATCTGCAGGCTCCAGTGGATGATGCTGCTTGAATGTAAGCTGA 584
QY 61 AAACAAACAGAGAGACTGTGTGTGTGTGGGAGATGTAATCATTCCTCCACAAC 120
DB 583 AAACAAACAGAGAGACTGTGTGTGTGTGGGAGATGTAATCATTCCTCCACAAC 524
QY 121 CTGC 124
DB 523 CTGC 520

RESULT 23
BU601136 678 bp mRNA linear EST 20-SEP-2002
LOCUS BU601136
DEFINITION AGENCOURT 10018944 NIH_MGC_142 Homo sapiens cDNA clone
IMAGE:6495009 5', mRNA sequence.
VERSION BU601136
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 678)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC <http://mgs.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LNCM2672 row: c column: 10
High quality sequence strop: 527.

FEATURES
source Location/Qualifiers
1..678

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6495009"
/tissue_type="mixed (pool of 40 RNAs)"

/lab host="DH10B (T1-phage-resistant)"
/clone lib="NIH MGC 142"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggccgctccggc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-AAGCAGTGTATCAACGAGGAGGAGTGGCTTATGATGCAAGCTGA
60
5'-ATTCTAGAGCGGAGCGGCGGCGGAGTGGCTTATGATGCAAGCTGA
60
Created library was constructed using the Clontech
Creator SMART kit and size-selected to contain the >0.5 kb
size fraction (other fractions present in NIH_MGC_141).
Library created in the laboratory of M. Brownstein (NIH,
NIH). Note: this is a NIH MGC library." 26 others

BASE COUNT 185 a 129 c 160 g 178 t

ORIGIN

Query Match 100.0%; Score 124; DB 13; Length 678;
Best Local Similarity 100.0%; Pred. No. 3.6e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATCGTGGCGCATCTGCGAGGCTCCAGGTGATGATGCTCTTATGATGCAAGCTGA 60
DB 13 CGATCGTGGCGCATCTGCGAGGCTCCAGGTGATGATGCTCTTATGATGCAAGCTGA 72

QY 61 AAACAACAAGAGAGAGCTGTGTGTGTGTGTGTGGGAGATGATATCAATTCCTTCCACACTG 120
DB 73 AAACAACAAGAGAGAGCTGTGTGTGTGTGTGTGGGAGATGATATCAATTCCTTCCACACTG 132

QY 121 CTGC 124
DB 133 CTGC 136

RESULT 24
LOCUS AI338342 702 bp mRNA linear EST 13-FEB-1999
DEFINITION q96g03.x1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
IMAGE:1939252.3 similar to WP:R10A10.2 CE12670 ZINC FINGER PROTEIN
; mRNA sequence.

ACCESSION AI338342
VERSION AI338342.1 GI:4075269
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 702)
NCI-CCRP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 854 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence scop: 455.
Location/Qualifiers

FEATURES
source 1..702
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1939252"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total fetus Nb2HF8 9w"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - 01190(dT) primer [5'
TGTTACCATCTGAGAGAGAGGCGGCGGCTTAATTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 189 a 175 c 137 g 201 t

ORIGIN

Query Match 100.0%; Score 124; DB 9; Length 702;
Best Local Similarity 100.0%; Pred. No. 3.6e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATCGTGGCGCATCTGCGAGGCTCCAGGTGATGATGCTCTTATGATGCAAGCTGA 60
DB 628 CGATCGTGGCGCATCTGCGAGGCTCCAGGTGATGATGCTCTTATGATGCAAGCTGA 569

QY 61 AAACAACAAGAGAGAGCTGTGTGTGTGTGTGTGGGAGATGATATCAATTCCTTCCACACTG 120
DB 568 AAACAACAAGAGAGAGCTGTGTGTGTGTGTGTGGGAGATGATATCAATTCCTTCCACACTG 509

QY 121 CTGC 124
DB 508 CTGC 505

RESULT 25
LOCUS BM704660 706 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-C11-af0-o-15-0-UI-r2 UI-E-C11 Homo sapiens cDNA clone
UI-E-C11-af0-o-15-0-UI 5', mRNA sequence.

ACCESSION BM704660
VERSION BM704660.1 GI:19017918
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 706)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
source 1..706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-af0-o-15-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

ACCESSION	BG766992
VERSION	BG766992.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	NIH-MGC http://mhc.nci.nih.gov/.
TITLE	1 (bases 1 to 754)
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished Contact: Robert Strausberg, Ph.D. Email: cga@sra-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTF cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Place: LNCMI743 row: e column: 04 High quality sequence stop: 750. Location/Qualifiers 1..754
FEATURES	/organism="Homo sapiens" /mol_type="RNA" /db_xref="taxon:9606" /clone="IMAGE:4870251" /tissue.type="melanotic melanoma, high MDR (cell line)" /lab_host="DH10B (phage-resistant)" /clone.lib="NIH MGC 49" /note="Organ: skin; Vector: pOTB7; Site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."
BASE COUNT	205 a 160 c 197 g 192 t
ORIGIN	" "
Query Match	100.0%; Score 124; DB 12; Length 754;
Best Local Similarity	100.0%; Pred. No. 3,7e-54;
Matches 124; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 CGATCGGCCGCATCTCAGGGTCAGGTATGATGCTGTCTTAATGTCAGCTGA 60
Db	129 CGATCGGCCGCATCTCAGGGTCAGGTATGATGCTGTCTTAATGTCAGCTGA 188
Oy	61 AAACAACAAAGAGAAGCTGTGTGTGGTGAGAGATGATCATTTCTTCCACAACCTG 120
Db	189 AAACAACAAAGAGAAGCTGTGTGTGTGGTGAGAGATGATCATTTCTTCCACAACCTG 248
Oy	121 CTGC 124
Db	249 CTGC 252
RESULT 31	
LOCUS	Bi601470
DEFINITION	Bi601470 767 bp mRNA linear EST 07-SEP-2001
ACCESSION	603249053r1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5300845 5',
VERSION	Bi601470
KEYWORDS	mRNA sequence.
SOURCE	Bi601470.1 GI:15494409
ORGANISM	EST.
	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 767)

AUTHORS		NIH-MGC http://mgc.nci.nih.gov/.
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL		Unpublished.
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgaabs-remail.nih.gov Tissue Procurement: Miklos Palcovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA library Arrayed by: The I.M.A.G.E Consortium (ULNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E Consortium/ULNL at: http://image.lnl.gov Plate: LHAM1761 row: j column: 14 High quality sequence stop: 701.
FEATURES		location/qualifiers
Source		1..767
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5300845"	
	/tissue_type="hypothalamus"	
	/lab_host="DH10B"	
	/clone_lib="NIH_MGC_96"	
	/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgag) ; Oligo-dt primed using primer 5'-TTTTTTTCTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to Rot 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."	
BASE COUNT	192 a	173 c 214 g 188 t
ORIGIN		
Query Match	100.0%; Score 124;	DB 12; Length 767;
Best Local Similarity	100.0%; Prid. NO. 3.7e-54;	
Matches 124;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
OY	1 CGATACGTGGCCCATCTGCAGAGGTCGCAAGTGATGATCGCTCTTAGATGTCAAAGCTGA 60	
Dd	175 CGATACGTGGCCCATCTGCAGAGGTCGCAAGTGATGATGATCGCTCTTAGATGTCAAAGCTGA 234	
OY	61 AAAAACAACAGAGGACTGTGTGTGTGTGGGAGAATTATATCTTCACAACACTG 120	
Dd	235 AAAAACAACAGAGGACTGTGTGTGTGTGGGAGAATTATATCTTCACAACACTG 294	
OY	121 CTGC 124	
Dd	295 CTGC 298	
RESULT 32		
LOCUS BG111145	793 bp	mRNA linear EST 30-JAN-2001
DEFINITION	60228155F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4369332 5'	
ACCESSION	BG111145	
VERSION	BG111145.1 GI:12604651	
KEYWORDS	mRNA sequence.	
SOURCE	EST.	
ORGANISM	Homo sapiens (human)	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE	1 (bases 1 to 793)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: life Technologies, Inc.	

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LLM10024 row: m column: 13
High quality sequence stop: 602.
Location/Qualifiers

FEATURES

source

1..793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4369332"
/issue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_86"
/note="Organ: bone; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.533 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 200 a 187 c 218 g 187 t 1 others
ORIGIN

Query Match 100.0%; Score 124; DB 10; Length 793;
Best Local Similarity 100.0%; Pred. No. 3.7e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTGGCCATCTGCAGGGTCCAGTGATGATGCTGCTTAGATGCAACTGA 60
DB 157 CGATACGTGGCCATCTGCAGGGTCCAGTGATGATGCTGCTTAGATGCAACTGA 216
QY 61 AAACAACAAGAGAGACTGTGTGTGTGGTGGGAGATGTAATCATCTCTCCACAAC 120
DB 217 AAACAACAAGAGAGACTGTGTGTGTGGTGGGAGATGTAATCATCTCTCCACAAC 276
QY 121 CTGC 124
DB 277 CTGC 280

RESULT 33
BI832643 797 bp mRNA linear EST 04-OCT-2001
LOCUS 603082021F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5221325 5',
DEFINITION mRNA sequence.
ACCESSION BI832643
VERSION BI832643.1 GI:15944193
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 797)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LLM1556 row: i column: 06
High quality sequence stop: 795.
Location/Qualifiers

FEATURES

source

1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:5221325"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT 208 a 179 c 212 g 198 t
ORIGIN

Query Match 100.0%; Score 124; DB 12; Length 797;
Best Local Similarity 100.0%; Pred. No. 3.8e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTGGCCATCTGCAGGGTCCAGTGATGATGCTGCTTAGATGCAACTGA 60
DB 171 CGATACGTGGCCATCTGCAGGGTCCAGTGATGATGCTGCTTAGATGCAACTGA 230
QY 61 AAACAACAAGAGAGACTGTGTGTGTGGTGGGAGATGTAATCATCTCTCCACAAC 120
DB 231 AAACAACAAGAGAGACTGTGTGTGTGGTGGGAGATGTAATCATCTCTCCACAAC 290
QY 121 CTGC 124
DB 291 CTGC 294

RESULT 34
BI858307 797 bp mRNA linear EST 10-OCT-2001
LOCUS 603384080F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5393131 5',
DEFINITION mRNA sequence.
ACCESSION BI858307
VERSION BI858307.1 GI:15999054
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 797)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LLM12001 row: o column: 20
High quality sequence stop: 718.
Location/Qualifiers

FEATURES

source

1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5393131"
/issue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_87"
/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN	BASE COUNT	229 a	166 c	204 g	198 t
Query Match		100.0%;	Score 124;	DB 12;	Length 797;
Best Local Similarity		100.0%;	Pred. No. 3.8e-54;		
Matches 124;	Conservative	0;	Mismatches	0;	Indels 0; Gaps 0
QY	1	CGATACGTGCGCCATCTGCAGGGTCCAGGTGATGATGCTGTCTTTAGATGTCAGACTGA	60		
Db	138	CGATACGTGCGCCATCTGCAGGGTCCAGGTGATGATGCTGTCTTTAGATGTCAGACTGA	197		
QY	61	AAACAAACAGAGAGACTGTGTGTGTGTGTGCGGGAGAGATATTCATTCTTCCACACTG	120		
Db	198	AAACAAACAGAGAGACTGTGTGTGTGTGTGCGGGAGAGATATTCATTCTTCCACACTG	257		
QY	121	CTGC 124			
Db	258	CTGC 261			

RESULT 35
BG708518

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

805 bp mRNA linear EST 07-MAY-2001
602670411n1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793091 5',
mRNA sequence.
BG708518
BG708518.1 GI:13985940
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini, Homnidae; Homo.
1 (bases 1 to 805)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraiki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10672 row: f column: 04
High quality sequence stop: 740.

```

FEATURES
source
1..805
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_image="493091"
/clone_type="hypothalamus"
/lab_host="DH10B"
/clone_id="NH_MGC_96"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgac
); Oligo-dt primed using primer 5'-TTTATTTTATTTTATTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Garinai, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NH_MGC library."
BASE COUNT
214 a 183 c 214 g 194 t
ORIGIN
Query Match 100.0%; Score 124; DB 10; Length 805;
Best Local Similarity 100.0%; Pred. No. 3.8e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

QY	1	CGATACGTCGCGCATCTGGAGGGTCCAGATGATGCCCTGTGTGATGATCAAGCTGA	60
Db	166	CGATACGTCGCGCATCTGGAGGGTCCAGATGATGCCCTGTGTGATGATCAAGCTGA	225
QY	61	AACCAACAAGAGACTGTGTGTGCTGGGAGAAATGATCATTCCTCCACAATCTG	120
Db	226	AACCAACAAGAGACTGTGTGTGCTGGGAGAAATGATCATTCCTCCACAATCTG	285
QY	121	CTGC	124
Db	286	CTGC	289

RESULT	36
B1457840	
LOCUS	
DEFINITION	B1457840 Homo sapiens CDNA clone IMAGE:5277723 5' , linear EST 21-AUG-2000
ACCESSION	U00901
VERSION	1.0
KEYWORDS	mRNA sequence.
SOURCE	B1457840 B1457840.1 GI:15248496
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 818)
TITLE	NH-MGC http://mgc.ncl.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished Contact: Robert Strausberg, Ph.D.

cdna Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LNLU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLU at:
<http://image.llnl.gov>
Plate: LLMML1701 row: g column: 04
High quality sequence stop: 780.
Location/Qualifiers
1. .818

	Query Match	Best Local Similarity	Matches 124; Conservative	Score 124; Pred. No. 3.8e-54; Mismatches 0; Indels 0; Gaps 0	DB 12; Length 818;
Oy	1	CGATAGTGGCCATCTGCAGGGGTCAGGTATGATGCCGTCTTGATGATGTCAGCTGA	60		
Db	157	CGATAGTGGCCATCTGCAGGGGTCAGGTATGATGATGCCGTCTTGATGATGTCAGCTGA	216		
Oy	61	AAACAAACAAGAGACTGTGTGTGGTCTGGGAGAAATGATATCATTTCTCCCAACTG	120		
Db	217	AAACAAACAAGAGACTGTGTGTGGTCTGGGAGAAATGATATCATTTCTCCCAACTG	276		

RESULT 37	121 CTGC 124
B1668735	
LOCUS	277 CTGC 280
DEFINITION	B1668735 822 bp mRNA linear EST 12-SEP-2001 603293174r1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5312226 5', mRNA sequence.
ACCESSION	B1668735
VERSION	B1668735.1
KEYWORDS	GI:15582968
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 822)
TITLE	NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished Contact: Robert Strausberg, Ph.D. Email: cga@bbs-rt@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLM11791 row: d column: 19 High quality sequence stop: 788. Location/Qualifiers 1. 822 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5312226" /tissue_type="hypothalamus" /lab_host="DH10B" /clone_id="NIH_MGC_96" /note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (grecgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT	219 a 182 c 219 g 202 t
ORIGIN	
Query Match	100.0%; Score 124; DB 12; Length 822;
Best Local Similarity	100.0%; Pred. No. 3.8e-54;
Matches 124; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CGATTCGTCGCCCATCTGCAAGGCTCAGAGTGATGATGCTGCTTAAATGTCAAGCTGA 60
Db	169 CGAATACGTCGCCCATCTGCAAGGCTCAGAGTGATGATGCTGCTTAAATGTCAAGCTGA 228
QY	61 AAACAAACAAGAGACTGTGTGTGTGTCGCGGAGATGATATCATCTCTCCACAACTGG 120
Db	229 AAACAAACAAGAGACTGTGTGTGTGCTGTGGGAGATGATATCATCTCTCCACAACTGG 288
QY	121 CTGC 124
Db	289 CTGC 282

LOCUS	BI828930	827 bp	mRNA	linear	EST 04-OCT-2001
DEFINITION	603075092p1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5166790 5', mRNA sequence.				
ACCESSION	BI828930				
VERSION	BI828930.1	GI:15940480			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 827)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgabbs-remail.nih.gov				
	Tissue Procurement: Life Technologies, Inc.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:				
	http://image.lnl.gov				
	Plate: LLM11414 row: h column: 23				
	High quality sequence start: 2				
	High quality sequence stop: 679.				
FEATURES	location/Qualifiers				
source	1..827				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:5166790"				
	/tissue_type="medulla"				
	/lab_host="DH10B"				
	/clone_id="NIH_MGC_119"				
	/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is Oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note this is a NIH_MGC Library."				
BASE COUNT	232 a 183 c 227 g 184 t				
ORIGIN	1 others				
	Query Match 100.0%; Score 124; DB 12; Length 827;				
	Best Local Similarity 100.0%; Prid. No. 3.8e-54;				
	Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY	1 CGATACGGCGGCACATCTGCAGCGGTCGAGGTGATGGATGCGTCTTGATGTCGAAGCTGA 60				
Db	153 CGATACGGCGGCACATCTGCAGCGGTCGAGGTGATGGATGCGTCTTGATGTCGAAGCTGA 212				
OY	61 AAACAAACAGAGAGACTGTGTGTGTGCTCTGGGAGAAATGATCATCTTCTCCACAACCTG 120				
Db	213 AAACAAACAGAGAGACTGTGTGTGTGCTCTGGGAGAAATGATCATCTTCTCCACAACCTG 272				
OY	121 CTGC 124				
Db	273 CTGC 276				
RESULT 39					
LOCUS	BG037022	840 bp	mRNA	linear	EST 24-JAN-2001
DEFINITION	60228734F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4374443 5', mRNA sequence.				
ACCESSION	BG037022				
VERSION	BG037022.1	GI:12432833			
KEYWORDS	EST.				

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eulacozoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 840)
AUTHORS	NIH-MGC http://mgs.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D.

FEATURES	Location/Qualifiers
source	1. .840

BASE COUNT	219 a	191 c	232 g	198 t
ORIGIN				

[illegible]

RESULT 40	
BG773490	847 bp mRNA linear EST 15-MAY-2006
LOCUS	BG773490
DEFINITION	602120212F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4637274 5' , mRNA sequence.
ACCESSION	BG773490
VERSION	BG773490
KEYWORDS	BG773490.1 GI:14084143
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 847)
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	

Unpublished
Journal
Comment
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shireki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LLAM10769 row: g column: 03
High quality sequence stop: 775.

BASE COUNT	199 a	205 c	242 g	201 t
ORIGIN				

QY	61	AAACAAACAGAGGACTGCTGTGTGATCTGGGGAGAAATGATCATTCCTCCAAACTG	120
Db	322	AAACAAACAGAGGACTGCTGTGTGATCTGGGGAGAAATGATCATTCCTCCAAACTG	381
QY	121	CTGC	124
Db	382	CTGC	385

RESULT 41				
B0876591				
LOCUS				
DEFINITION	B0876591	853 bp	mRNA	linear
	AGENCOCURT 8584244	lupski sympathetic_trunk	Homo sapiens	cDNA clone
	IMAGE:6192819	5'		mRNA sequence.
ACCESSION	B0876591			
VERSION	B0876591.1	GI:22268599		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.			
REFERENCE	1 (bases 1 to 853)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-x@mail.nih.gov Tissue Procurement: Dr. James R. Lupski cDNA Library Preparation: Life Technologies, Inc. cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILLUM) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be			

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov/
Plate: L13595 row: h column: 04
High quality sequence stop: 608.
Location/Qualifiers

FEATURES
source

1. 853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6192819"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski sympathetic trunk"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGGCTCG-3' and
5'-GACTAGTCTAGATCGCAGCGCCGCTT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."
BASE COUNT 231 a 191 c 225 g 205 t 1 others
ORIGIN

Query Match 100.0%; Score 124; DB 13; Length 853;
Best Local Similarity 100.0%; Pred. No. 3.8e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGATACGTGGCCCATCTGCAGGGTCCAGTGATGCTCTTAGATGCAAGCTGA 60
DB 191 CGATACGTGGCCCATCTGCAGGGTCCAGTGATGCTCTTAGATGCAAGCTGA 250
OY 61 AAACAAACAGAGAGAGCTGTGTGTGTGTGTGGAGATGTATCATTCCTCCACAACTG 120
DB 251 AAACAAACAGAGAGAGCTGTGTGTGTGTGTGGAGATGTATCATTCCTCCACAACTG 310
OY 121 CTGC 124
DB 311 CTGC 314

RESULT 42
BE747000 855 bp mRNA linear EST 15-SEP-2000
LOCUS 601580743F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3929459 5',
DEFINITION mRNA sequence.
ACCESSION BE747000.1 GI:10160992
VERSION BE747000.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homimidae; Homo.
REFERENCE 1 (bases 1 to 855)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L13595 row: e column: 12
High quality sequence stop: 767.
Location/Qualifiers

FEATURES
source

1. 855

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3929459"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 250 a 161 c 238 g 205 t 1 others
ORIGIN

Query Match 100.0%; Score 124; DB 10; Length 855;
Best Local Similarity 100.0%; Pred. No. 3.8e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGATACGTGGCCCATCTGCAGGGTCCAGTGATGCTCTTAGATGCAAGCTGA 60
DB 112 CGATACGTGGCCCATCTGCAGGGTCCAGTGATGCTCTTAGATGCAAGCTGA 171
OY 61 AAACAAACAGAGAGAGCTGTGTGTGTGTGTGGAGATGTATCATTCCTCCACAACTG 120
DB 172 AAACAAACAGAGAGAGCTGTGTGTGTGTGTGGAGATGTATCATTCCTCCACAACTG 231
OY 121 CTGC 124
DB 232 CTGC 235

RESULT 43
BG753323 856 bp mRNA linear EST 15-MAY-2001
LOCUS 602731740F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4875489 5',
DEFINITION mRNA sequence.
ACCESSION BG753323
VERSION BG753323.1 GI:14063976
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homimidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L13595 row: o column: 10
High quality sequence stop: 767.
Location/Qualifiers

FEATURES
source

1. 856
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4875489"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
OY	61	AAACAAACAGAGGACTGCTGTGTGGTCTCGGGGAGAAATGATTCCTTCACAACTG	CB991438	886 bp	mRNA	linear	EST 01-MAY-2001					
DB	276	AAACAAACAGAGGACTGCTGTGTGGTCTCGGGGAGAAATGATTCCTTCACAACTG	AGNCOURT_13621565	NIH MGC_148	Homo sapiens	cDNA clone						
OY	121	CTGC 124	CB991438	1	GI:30285958							
DB	336	CTGC 339	CB991438	1	GI:30285958							
OY	121	CTGC 124	CB991438	1	GI:30285958							
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	61	AAACAAACAGAGGACTGCTGTGTGGTCTCGGGGAGAAATGATTCCTTCACAACTG	CB991438	886 bp	mRNA	linear	EST 01-MAY-2001					
DB	231	AAACAAACAGAGGACTGCTGTGTGGTCTCGGGGAGAAATGATTCCTTCACAACTG	AGNCOURT_13621565	NIH MGC_148	Homo sapiens	cDNA clone						
OY	121	CTGC 124	CB991438	1	GI:30285958							
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	61	AAACAAACAGAGGACTGCTGTGTGGTCTCGGGGAGAAATGATTCCTTCACAACTG	CB991438	886 bp	mRNA	linear	EST 01-MAY-2001					
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	121	CTGC 124	CB991438	1	GI:30285958							
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	61	AAACAAACAGAGGACTGCTGTGTGGTCTCGGGGAGAAATGATTCCTTCACAACTG	CB991438	886 bp	mRNA	linear	EST 01-MAY-2001					
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	121	CTGC 124	CB991438	1	GI:30285958							
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	61	AAACAAACAGAGGACTGCTGTGTGGTCTCGGGGAGAAATGATTCCTTCACAACTG	CB991438	886 bp	mRNA	linear	EST 01-MAY-2001					
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	121	CTGC 124	CB991438	1	GI:30285958							
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	61	AAACAAACAGAGGACTGCTGTGTGGTCTCGGGGAGAAATGATTCCTTCACAACTG	CB991438	886 bp	mRNA	linear	EST 01-MAY-2001					
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	121	CTGC 124	CB991438	1	GI:30285958							
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	61	AAACAAACAGAGGACTGCTGTGTGGTCTCGGGGAGAAATGATTCCTTCACAACTG	CB991438	886 bp	mRNA	linear	EST 01-MAY-2001					
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290								
OY	121	CTGC 124	CB991438	1	GI:30285958							
DB	231	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	CGATACGTGCGGCATCTGACAGGCTCCAGGTGATGATGCTGCTCTTAAGTGTCAAGCTTA	290	</							

[illegible]

REFERENCE 1 (bases 1 to 918)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM11457 row: m column: 20
High quality sequence stop: 732.
Location/Qualifiers

FEATURES
source 1..918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5183419"
/lab_host="DH10B"
/clone_lib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC library."

BASE COUNT 269 a 196 c 247 g 206 t
ORIGIN

Query Match 100.0%; Score 124; DB 12; Length 918;
Best Local Similarity 100.0%; Pred. No. 3.9e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGCGCCATCTGCAGGCTCAGATGATGATCCCTCTTAATGTCGAAGCTGA 60
DB 152 CGATACGCGCCATCTGCAGGCTCAGATGATGATCCCTCTTAATGTCGAAGCTGA 211
QY 61 AAACAACAAGAGACTGTGTGCTGGGAGATGATCATCTTCCTCCCAACTG 120
DB 212 AAACAACAAGAGACTGTGTGCTGGGAGATGATCATCTTCCTCCCAACTG 271
QY 121 CTGC 124
DB 272 CTGC 275

RESULT 49
BUI92094 922 bp mRNA linear EST 04-SEP-2002
LOCUS AGENCOURT 7968580 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6011203
DEFINITION 5', mRNA sequence.
ACCESSION BUI92094
VERSION BUI92094.1 GI:22706078
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 922)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM13200 row: h column: 20
High quality sequence stop: 570.
Location/Qualifiers

FEATURES
source 1..922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6011203"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_68"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."

BASE COUNT 249 a 202 c 250 g 220 t 1 others
ORIGIN

Query Match 100.0%; Score 124; DB 13; Length 922;
Best Local Similarity 100.0%; Pred. No. 3.9e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGCGCCATCTGCAGGCTCAGATGATGATCCCTCTTAATGTCGAAGCTGA 60
DB 155 CGATACGCGCCATCTGCAGGCTCAGATGATGATCCCTCTTAATGTCGAAGCTGA 214
QY 61 AAACAACAAGAGACTGTGTGCTGGGAGATGATCATCTTCCTCCCAACTG 120
DB 215 AAACAACAAGAGACTGTGTGCTGGGAGATGATCATCTTCCTCCCAACTG 274
QY 121 CTGC 124
DB 275 CTGC 278

RESULT 50
BG111792 947 bp mRNA linear EST 30-JAN-2001
LOCUS 602285379F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4372819 5',
DEFINITION mRNA sequence.
ACCESSION BG111792
VERSION BG111792.1 GI:12605298
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 947)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM10033 row: n column: 20
High quality sequence stop: 682.
Location/Qualifiers

FEATURES
source 1..947
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4372819"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_86"
/note="Organ: bone; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.53 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library." 1 others

BASE COUNT 228 a 241 c 267 g 210 t 1 others
ORIGIN

Query Match 100.0%; Score 124; DB 10; Length 947;
Best Local Similarity 100.0%; Pred. No. 3.9e-54;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGATACGTGCGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGTCAGCTGA	60
Db	194	CGATACGTGCGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGTCAGCTGA	253
Qy	61	AAACAAACAGAGACTGTGTTGTGCTGCGGAGATGTATCATTCCTCCACACTG	120
Db	254	AAACAAACAGAGACTGTGTTGTGCTGCGGAGATGTATCATTCCTCCACACTG	313
Qy	121	CTGC	124
Db	314	CTGC	317

Search completed: November 7, 2003, 11:54:01
Job time : 1272.21 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 11:59:37 ; Search time 2874.08 Seconds
(without alignments)
9640.351 Million cell updates/sec

Title: US-09-509-779-1

Perfect score: 1140
Sequence: 1 GTTCTGGCGCGCGCCATG.....ACTAATTCATCAATTAATG 1140

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_hiv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1069.4	93.8	1132	11	AK003963 Mus muscu
2	1041	91.3	1114	11	AK011328 Mus muscu
3	1032.4	90.6	1084	11	AK003248 Mus muscu
4	1004.4	88.1	1078	11	AK002628 Mus muscu

5	875	76.8	997	14	BY704214	BY704214
6	856.2	75.1	948	14	BY703542	BY703542
7	836.2	73.4	989	13	BH840985	BH840985
8	833.6	73.1	901	14	CB204521	CB204521
9	802.4	70.4	1185	13	BQ964159	BQ964159
10	781.8	68.6	968	13	BQ964115	BQ964115
11	765	67.1	820	12	BT732003	BT732003
12	753.4	66.1	1079	12	BT140936	BT140936
13	740	64.9	882	14	CA980368	CA980368
14	739.6	64.9	938	14	CB209388	CB209388
15	732.2	64.2	965	14	BY710626	BY710626
16	725.4	63.6	743	14	CA323814	CA323814
17	719.4	63.1	970	13	BUS13948	BUS13948
18	707.2	62.0	728	13	BO571112	BO571112
19	700.8	61.5	722	11	AK007588	AK007588
20	687.2	60.3	721	14	BY707632	BY707632
21	686.6	60.2	731	12	BI647378	BI647378
22	682.8	59.9	889	14	CB196142	CB196142
23	680.2	59.7	867	14	CA464071	CA464071
24	673.4	59.1	978	12	BT156184	BT156184
25	671	58.9	787	14	CB318729	CB318729
26	665.4	58.4	916	12	BS917948	BS917948
27	664.6	58.3	944	14	CA977674	CA977674
28	653.4	57.3	676	13	BQ177475	BQ177475
29	651.6	57.2	686	12	BQ044727	BQ044727
30	651.4	57.1	754	12	BT147544	BT147544
31	651	57.1	704	10	BF018804	BF018804
32	649.2	56.9	946	10	BF83735	BF83735
33	647.8	56.8	767	14	CB571468	CB571468
34	646.2	56.7	737	12	BT556918	BT556918
35	629	55.2	642	14	BY737338	BY737338
36	627.2	55.0	920	14	CB209237	CB209237
37	624.4	54.8	822	10	BF302267	BF302267
38	612.2	53.7	755	10	BF579825	BF579825
39	597.2	52.4	620	10	BE628111	BE628111
40	596	52.3	651	9	AM555800	AM555800
41	592.4	52.0	728	14	CB951114	CB951114
42	586.8	51.5	653	14	BY757681	BY757681
43	584	51.2	624	12	BM946344	BM946344
44	582.8	51.1	757	14	CB318426	CB318426
45	580.4	50.9	779	14	CD241479	CD241479

ALIGNMENTS

RESULT 1
AK003963
LOCUS
DEFINITION Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:110029M05 product:ring finger protein 7, full insert sequence.
ACCESSION AK003963
VERSION AK003963.1 GI:12834939
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
PUBLISHED	20030913
REFERENCE AUTHORS	4 Kawai, J., Shigaawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nialaid, I., Pesole, G., Quackenbush, J., Schirml, L.M., Stabli, F., Suzuki, R., Tomita, M., Wagner, T., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamaya, M., Lee, N.H., Lyons, P., Ringdon, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ritch, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyokawa, K., Wang, K.H., Weitz, C., Whitlaker, C., Wilming, L., Wynshaw-Borie, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
PUBLISHED	21085660
REFERENCE AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE AUTHORS	6 (bases 1 to 1132) Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Aikawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hara, A., Hatanaka, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shigaawa, A., Shiraki, T., Sogabe, Y., Suwaki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyata, Y., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Morimatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGGAGAGCGCCGCACTGACTGATTTTCTTTTCTTTCN 3'], cDNA was prepared by using triethanolamine thermo-activated reverse transcriptase

FEATURES									
source	Location/Qualifiers	Xhol. Host: SOLR.							
	1. .1132								
	/organism="Mus musculus"								
	/mol_type="mRNA"								
	/strain="C57BL/6J"								
	/db_xref="PANTOM,DB:1110029W05"								
	/db_xref="MGI:1892527"								
	/db_xref="taxon:10090"								
	/clone="1110029W05"								
	/issue_type="whole body"								
	/clone_lib="RIKEN full-length"								
	/dev_stage="16-day embryo"								
	15..355								
	/note="putative								
	ring finger protein 7 (MGD MGI:1337096)"								
	/db_xref="MGI:1337096"								
	1113. .1118								
	/note="putative"								
	1132								
polyA_site									
	/note="putative"								
BASE COUNT	298 a 224 c 286 g 324 t								
ORIGIN									
Query Match	93.8%; Score 1069.4; DB 11; Length 1132;								
Best Local Similarity	99.3%; Pred. No. 1.4e-218;								
Matches 1127; Conservative	0; Mismatches 1; Indels 7; Gaps 5								
QY	4 CTGCGCGCGCCATGTGCGCGAGCGTGGAGAGACGCGAGGAACCTTCGTCCTTTCTTGCA	63							
Db	2 CTGCGCGCGCGCCATGTGCGCGAGCGTGGAGAGACGCGAGGAACCTTCGTCCTTTCTTGCA	61							
QY	64 CTCGGGAGACGAGCGCTCCAGTCCGAGGAGGAGACAAAGTGTCTCTCAAGAAAGTGGAA	1233							
Db	62 CTCGGGAGACGAGCGCTCCAGTCCGAGGAGGAGACAAAGTGTCTCTCTCAAGAAAGTGGAA	120							
QY	124 CCGCGTAGCCATGTGAGAGTGAGAGCTTGAGCTGATACCTGTGCATCTGACAGGGTCCA	1833							
Db	121 CCGCGTAGCCATGTGAGAGTGAGAGCTTGAGCTGATACCTGTGCATCTGACAGGGTCCA	180							
QY	184 GGTGATGATGCTCTGCTCTTCGATGTCAAGCTGAAAAAACAAGACAGAGACTGTGTGGGT	2433							
Db	181 GGTGATGATGCTCTGCTCTTCGATGTCAAGCTGAAAAAACAAGACAGAGACTGTGTGGGT	240							
QY	244 CTGGGAGAGGTAAACCAATCCCTTCACAACTGCTGCATGCTCCCTGGGAGTGAACAGAA	303							
Db	241 CTGGGAGAGGTAAACCAATCCCTTCACAACTGCTGCATGCTCCCTGGGAGTGAACAGAA	300							
QY	304 CAATGCTGCTCTGTGTGCCACAGACCTGGTATGTCMAAAGATGGCAATGAGAGGT	3633							
Db	301 CAATGCTGCTCTGTGTGCCACAGACCTGGTATGTCMAAAGATGGCAATGAGAGGT	360							
QY	364 GGGCCAGAGCGCTCTGTGTGTGTTGCTGACCTCTGACAAAGACCTAAACACTGACGGGAT	4233							
Db	361 GGGCCAGAGCGCTCTGTGTGTGTTGCTGACCTCTGACAAAGACCTAAACACTGACGGGAT	420							
QY	424 TCATCTTGAGAGAGAGAGATGCTGTGCGCTTTGAGACTCAACAAAGCTTGCTTTAT	4833							
Db	421 TCATCTTGAGAGAGAGAGATGCTGTGCGCTTTGAGACTCAACAAAGCTTGCTTTAT	480							
QY	484 TAAATTTGCTGTTTACTTTGGAAATTTCTTACAAATTAAGATPATTTGTTAAAAATGCG	5433							
Db	481 TAAATTTGCTGTTTACTTTGGAAATTTCTTACAAATTAAGATPATTTGTTAAAAATGCG	540							
QY	544 CTTTCCACCTCTGAGTGTGTGTGTGTATGAATGATGTAAGAGAGGAGAACACACAGAA	6033							
Db	541 CTTTCCACCTCTGAGTGTGTGTGTGTATGAATGATGTAAGAGAGGAGAACACACAGAA	600							
QY	604 AATGATCTTTGTTTATCTGTACCCAGACT--GGAACATTGTGTTCAGAAAGAACATT	6600							

Db	Accession	Version	KeyWords	Source	Organism	Reference	Authors	Title	Journal	Medline	Pubmed
Db	601	AK011328									
Qy	661	AK011328									
Db	661	AK011328									
Qy	721	AK011328									
Db	721	AK011328									
Qy	781	AK011328									
Db	781	AK011328									
Qy	841	AK011328									
Db	841	AK011328									
Qy	901	AK011328									
Db	901	AK011328									
Qy	961	AK011328									
Db	961	AK011328									
Qy	1021	AK011328									
Db	1021	AK011328									
Qy	1081	AK011328									
Db	1081	AK011328									
Qy	1141	AK011328									
Db	1141	AK011328									
Qy	1201	AK011328									
Db	1201	AK011328									
Qy	1261	AK011328									
Db	1261	AK011328									
Qy	1321	AK011328									
Db	1321	AK011328									
Qy	1381	AK011328									
Db	1381	AK011328									
Qy	1441	AK011328									
Db	1441	AK011328									
Qy	1501	AK011328									
Db	1501	AK011328									
Qy	1561	AK011328									
Db	1561	AK011328									
Qy	1621	AK011328									
Db	1621	AK011328									
Qy	1681	AK011328									
Db	1681	AK011328									
Qy	1741	AK011328									
Db	1741	AK011328									
Qy	1801	AK011328									
Db	1801	AK011328									
Qy	1861	AK011328									
Db	1861	AK011328									
Qy	1921	AK011328									
Db	1921	AK011328									
Qy	1981	AK011328									
Db	1981	AK011328									
Qy	2041	AK011328									
Db	2041	AK011328									

JOURNAL MEDLINE PUBMED REFERENCE	Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861 4
AUTHORS	Kawai,J., Shinaawa,A., Shibata,K., Yoshino,M., Itoh,M., Ichii,Y., Arikawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,T., Fukuoka,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamataka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadote,K., Matsuda,H., Ashburner,M., Batalov,S., Casavola,H., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pezole,G., Quackenbush,J., Schriml,L.M., Staab,D.I., Suzuki,R., Tomita,M., Wagner,L., Mashio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarrelli,R., Barsh,G., Blake,Y., Boftelli,D., Boljuga,N., Carrincci,P., de Bonaldo,M.F., Brownstein,M.J., Ball,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustinchik,S., Hill,D., Hottmann,M., Hume,D.A., Kamitani,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Monbaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Saeki,H., Satou,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-Oka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L., Wyshaw-Sorls-A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohseuki,S. and Hayashizaki,Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	5
AUTHORS	The FANTOM consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length CDNA's
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 114)
AUTHORS	Adachi,T., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arawaka,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Funuro,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hisakawa,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Katsukawa,T., Katoh,H., Kawai,J., Koijima,Y., Komoto,H., Koude,M., Koya,S., Kuwahara,C., Matsumura,T., Miyazaki,A., Nishii,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Shibata,K., Shibata,Y., Shinaawa,A., Shiraki,T., Sogabe,Y., Suwaki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suicho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, url:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGACGATCCAGACGACTCTTTTCTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTTCGAGTTTAATTAAATTAATCCCCCCCC 3']. cDNA was cleaved with XhoI and SclI. Cloning sites, 5' end: XhoI; 3' end: SclI. Host: SOLR.
FEATURES	location/Qualifiers
SOURCE	1..114 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J"

Query Match	Best Local Similarity	Matches 1109; Conservative	91.3%; Score 1041; Pred. No. 1.7e-212; 0; Mismatches 0; Indels 8; Gaps 6	DB 11; Length 1114;
Db	4	CTGGCGCGCCGCATGGCCGAGCGTGGAGAGCGGCGAGGAAACCTGCGTCTTTCTTCGCA	63	
Db	2	CTGGCGCGCCGCATGGCCGAGCGTGGAGAGCGGCGAGGAAACCTGCGTCTTTCTTCGCA	60	
Qy	64	CTTCGGGAGCGCAGGCTCCAACTCGGGAGGCGCAAGAATGTTCTCTCCAAAGTGGAA	123	
Db	61	CTTCGGGAGCGC-GGCTCCAACTCGGGAGGCGCAAGAATGTTCTCTCCAAAGTGGAA	119	
Qy	124	CGCGGTACCAATGTTGAGCTGGGAGCGTTGAGTGCATACCTGTGCCATCTGCAGGGTCCA	183	
Db	120	CGCGGTACCAATGTTGAGCTGGGAGCGTTGAGTGCATACCTGTGCCATCTGCAGGGTCCA	179	
Qy	184	GGTATGATGATGCTCTGCTTCATATGTCACCTGTAACAAACGAAGAGACTGTGTGGT	243	
Db	180	GGTATGATGATGCTCTGCTTCATATGTCACCTGTAACAAACGAAGAGACTGTGTGGT	239	
Qy	244	CTGGGAGAGGTAAACCAATCTTCCCAACAAGTCTGCATGTCCTGTGGTGAACAGAA	303	
Db	240	CTGGGAGAGGTAAACCAATCTTCCCAACAAGTCTGCATGTCCTGTGGTGAACAGAA	299	
Qy	304	CAATGCTGACCTCTGTGTGCGACGACGACTGGGTAGTCCAAAGATGCGAAATGAGAGGT	363	
Db	300	CAATGCTGACCTCTGTGTGCGACGACGACTGGGTAGTCCAAAGATGCGAAATGAGAGGT	359	
Qy	364	GGCCGAGGGGCTCCGTGTGTGTGTGTCACCTGGGCAAAAGCTAAACCTGAGGGGAT	423	
Db	360	GGCCGAGGGGCTCCGTGTGTGTGTGTCACCTGGGCAAAAGCTAAACCTGAGGGGAT	419	
Qy	424	TCATCCTTGAAGAGAGAGATGCTGTGCGCCTTTGAGACTCAACAAAGCTTCTTAT	483	
Db	420	TCATCCTTGAAGAGAGAGATGCTGTGCGCCTTTGAGACTCAACAAAGCTTCTTAT	479	
Qy	484	TAAATTTGTCTTTTACTTTTGGGAAATTCCTTCAATTTAAAGTAATTTGTAAATAATGCG	543	
Db	480	TAAATTTGTCTTTTACTTTTGGGAAATTCCTTCAATTTAAAGTAATTTGTAAATAATGCG	539	
Qy	544	CTTTTCTTACCTCTGGGTGTGTGTGTGATACGATTCATAGAAGACCGAACCAGAA	603	
Db	540	CTTTTCTTACCTCTGGGTGTGTGTGTGATACGATTCATAGAAGACCGAACCAGAA	599	
Qy	604	AATGATCTTTGTTTATCTGTACCACGACT--GGAACATGTGTTCACAAAGAACATT	660	
Db	600	AATGATCTTTGTTTATCTGTACCACGACTGTGGGAAACATGTGTTCACAAAGAACATT	659	
Qy	661	GTTTGATTATGCTTGAAGGTTAAATAATAGTAAACGAATGTTACAGTAAACAAATTA	720	
Db	660	GTTTGATTATGCTTGAAGGTTAAATAATAGTAAACGAATGTTACAGTAAACAAATTA	719	
Qy	721	ATGCAATGAAAAGCCGACTCTTCTTAATCTTTTGTGTGGAGAGGCAACGAGGC	780	
Db	720	ATGCAATGAAAAGCCGACTCTTCTTAATCTTTTGTGTGGAGAGGCAACGAGGC	779	
Qy	781	CACCTGCTGTCTTCACTTGTCTGTGATAGAGATTTTAACCTGCACTCACTGAAGGCG	840	
Db	780	CACCTGCTGTCTTCACTTGTCTGTGATAGAGATTTTAACCTGCACTCACTGAAGGCG	839	

QY	841	TAACGTCCGGTAAACGTAAATGAATGAGGTACGTGCGGTAAACCGCTTTGTCCTCGAC	900
Db	840	TAACTGTGGGTAAACTGTATATATATGCGTAACTGTGGGTAAACGCTTTGTCCTCGAC	899
QY	901	TTCTCCATCTCTTGAAGCTTGGCCAGGAAACCTGGATTTGTTCAACCACTTAGTTCTTAAAGAAC	960
Db	900	TTCTCCATCTTGAAGCTTGGCCAGGAAACCTGGATTTGTTCAACCACTTAGTTCTTAAAGAAC	959
QY	961	TGTTTTCTGTTTTTGGCCGAAGGTGTATTTGATGTTTTTACGTCAAAATATATTAGAGAAA	1020
Db	960	TGTTTTCTGTTTTTGG-CGAAGGTGTATTTGATGTTTTTACGT-AAAATATTAGAGAAA	1017
QY	1021	ATGGCTACTAGTAACTAAACACTGAAGTTCATTATGCAATGTTTTTAATAAATATTGTGCTT	1080
Db	1018	AT-GCTTACGTAGTAACTAAACACTGAAGTTCATTATGCAATGTTTTTAATAAATATTGTGCTT	1076
QY	1081	TGAGTTATTTAAAGTTTGATATATATATACCTTTAAATCAT	1117
Db	1077	TGAGTTATTTAAAGTTTGATATATATATACCTTTAAATCAT	1113
RESULT 3			
AK003248		1084 bp	linear
LOCUS			
DEFINITION	AK003248	Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:1110001M08 product:ring finger protein 7, full insert sequence.	
ACCESSION	AK003248		
VERSION	AK003248.1	GI:12833795	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	1	Carninci, P. and Hayashizaki, Y.	
TITLE		High-efficiency full-length cDNA cloning	
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE		99279253	
PUBMED		10349636	
REFERENCE			
AUTHORS	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.	
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)	
MEDLINE		20499374	
PUBMED		11042159	
REFERENCE			
AUTHORS	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, T., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.	
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer	
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)	
MEDLINE		20530913	
PUBMED		11076861	
REFERENCE			
AUTHORS	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aikawa, T., Hara, A., Fukunishi, Y., Komoto, H., Adachi, D., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaoka, I., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Koehli, A., Kuehl, P., Lewis, S., Matsuo, Y., Nakado, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balarelli, R., Barish, G., Blake, J., Boffelli, D., Bojunga, N.,	


```

QY 901 TTCTCCATCTTTGATGCTGGCCGAGAACCTGGATTGTTCAACCACTTAGTCTTAAGAAG 960
DB 902 TTCTCCATCTTTGATGCTGGCCGAGAACCTGGATTGTTCAACCACTTAGTCTTAAGAAG 961
QY 961 TGTTCCTGCTTTTGGCCGAGAGTGTATGTTATGTTTAAAGAAATTTAGTGGAAA 1020
DB 962 TGTTCCTGCTTTTGGCCGAGAGTGTATGTTTAAAGTTC-AAAATTTAGTGGAAA 1019
QY 1021 ATGGCTTAGTATGATACCTGAAGTCAATGATGCAATGTTTAAATAATTTGCTT 1080
DB 1020 AT-GCTTAGTATGATACCTGAAGTCAATGATGCAATGTTTAAATAATTTGCTT 1078
QY 1081 TGAGTT 1086
DB 1079 TGAGCT 1084

RESULT 4
AC002628
LOCUS
DEFINITION
Mus musculus adult male kidney cDNA, RIKEN full-length enriched
library, clone:0610013021 product:ring finger protein 7, full
insert sequence.
AC002628.1 GI:12832752
VERSION
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
10349636
PUBMED

REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, N., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
11042159
PUBMED

REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
11076861
PUBMED

JOURNAL
PUBMED
REFERENCE
AUTHORS

TITLE
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
Saito, T., Okazaki, Y., Goto, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Niki, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsi, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. U., Bulc, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzei, J., Mombarts, P., Nordone, P.,

```

```

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

FEATURES
source
1.1078
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:0610013021"
/db_xref="MGI:1893044"
/db_xref="taxon:10090"
/clone="0610013021"
/sex="male"
/tissue_type="kidney"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
15.354
/note="putative
ring finger protein 7 (MGI:1337096)"
1060.1065
/note="putative"
1078
/note="putative"
BASE COUNT 276 a 216 c 283 g 303 t

COMMENT
further details
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGCGCGCCGACGCTGAGTTTCTTTTCTTTTCTT 3']. cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence [5'
GAGAGAGAGAGAGCGCGCCGACGCTGAGTTTCTTTTCTTTTCTT 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI, 3' end:
XhoI. Host: SOLR.
URL: http://genome.gsc.riken.go.jp/ for
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGCGCGCCGACGCTGAGTTTCTTTTCTTTTCTT 3']. cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence [5'
GAGAGAGAGAGAGCGCGCCGACGCTGAGTTTCTTTTCTTTTCTT 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI, 3' end:
XhoI. Host: SOLR.

```


Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source 1. 997
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="1110029M05"
 /tissue_type="whole body"
 /dev_stage="18-day embryo"
 /clone_lib="RIKEN full-length enriched, 18-day embryo whole body"

BASE COUNT 250 a 206 c 270 g 271 t

ORIGIN

Query Match

Best Local Similarity 98.8%; Score 875; DB 14; Length 997;
 Matches 924; Conservative 0; Mismatches 5; Indels 6; Gaps 4;

11 GCCGCATGCGCCGACGTGAGAGACGCGAGAACCTCGCTCTTCTTCGACCTCCGG 70
 1 GCCGCATGCGCCGACGTGAGAGACGCGAGAACCTCGCTCTTCTTCGACCTCCGG 60
 71 AGCGCAGGCTCCAGTCCGAGAGCGCAAGATGTTCTCTCAAGAAGTGGAAACCGGTA 130
 61 AGCGCA-GCTCCAGTCCGAGAGCGCAAGATGTTCTCTCAAGAAGTGGAAACCGGTA 119
 131 GCCATGTGAGAGTGGAGAGCTAGTGCATCTGCGCATCGCAGGGTCCAGGTGATG 190
 120 GCCATGTGAGAGTGGAGAGCTAGTGCATCTGCGCATCGCAGGGTCCAGGTGATG 179
 191 GATGCTGCTTCGATGTCAAGCTGCAAAACAAGAGAGCTGTGTTGTTGGTGGGA 250
 180 GATGCTGCTTCGATGTCAAGCTGCAAAACAAGAGAGCTGTGTTGTTGGTGGGA 239
 251 GAGTGAACCATTCCTTCCACAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 310
 240 GAGTGAACCATTCCTTCCACAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
 311 TSCCCTCTGTCAGAGAGAGCTGGGTAGTCCAAAGATGCGCAATGAGAGTGGCCAG 370
 300 TSCCCTCTGTCAGAGAGAGCTGGGTAGTCCAAAGATGCGCAATGAGAGTGGCCAG 359
 371 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 430
 360 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
 431 TGAAGAGAGAGAGAGTGTGGGCTTTGAGACTACCCAAAGAGCTTTGTTTAAATTTG 490
 420 TGAAGAGAGAGAGAGTGTGGGCTTTGAGACTACCCAAAGAGCTTTGTTTAAATTTG 479
 491 TCTGTTTAAATTTGGAATTTCTCAATTAAGATTAATTTGTTAAATTTGCTTTCT 550
 480 TCTGTTTAAATTTGGAATTTCTCAATTAAGATTAATTTGTTAAATTTGCTTTCT 539
 551 ACCCTGT 610
 540 ACCCTGT 599

QY 611 TTGTTATCTGTACCCACGAC--TGAACATTTGTTTACACAGAGAACTGTTGTG 667
 Db 600 TTTGTTATCTGTACCCACGACCTTTGGAACATTTGTTTACAGAGAACTTTGTTGTG 659
 QY 668 TTTATGCTGAGGGTTAAAAATAGATTAAGATTTAGAGTAACTAAATATGACTT 727
 Db 660 TTTATGCTGAGGGTTAAAAATAGATTAAGATTTAGAGTAACTAAATATGACTT 719
 QY 728 GAAAGCGCACTCTCTTAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 787
 Db 720 GAAAGCGCACTCTCTTAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 779
 QY 788 CTGCTCTTATTTGCTGTGAATGAGAGATTTTAACTGCACTAGTGAAGAGCGTACTGT 847
 Db 780 CTGCTCTTATTTGCTGTGAATGAGAGATTTTAA-CTGCACTAGTGAAGAGCGTAACTGT 838
 QY 848 CGGGTAACTGTATATATGCGGTAACTGTGCGGTAAAGCGCTTTGCTCTGACTTCTCA 907
 Db 839 CGGGTAACTGTATATATGCGGTAACTGTGCGGTAAAGCGCTTTGCTCTGACTTCTCA 898
 QY 908 TCTTTGACTTGGCCAGGAGAGCGCTGATTTGTTCAAC 942
 Db 899 TCTTTGACTTGGCCA-GAAGGCTGATTTGTTCAAC 932

RESULT 6

BY703542

LOCUS

DEFINITION

BY703542 RIKEN full-length enriched, 18-day embryo whole body Mus musculus cDNA clone 1110001M08 5', mRNA sequence.

ACCESSION

BY703542

VERSION

BY703542.1 GI:27114651

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

TITLE

JOHNNAL MEDLINE 22354683

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sueniro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sec.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, M., Hayashida, K., Hirozane, T., Hori, F., Imetani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koyama, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichipillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

```

1..988
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1110001M08"
/feature_type="whole body"
/dev_stage="18-day embryo"
/clone_lib="RIKEN full-length enriched, 18-day embryo whole body"

BASE COUNT      247 a      205 c      271 g      265 t
ORIGIN
Query Match      75.1%; Score 856.4; DB 14; Length 988;
Best Local Similarity 98.7%; Pred. No. 5.7e-173;
Matches 906; Conservative 0; Mismatches 6; Indels 6; Gaps 4;

QY 11 GCGCCATGGCCGACCTGGAGGACCGCGAGGAAACCTGGCTCTTTCTTCCGACTCCGGG 70
    |||
DB 1 GCGCCATGGCCGACCTGGAGGACCGCGAGGAAACCTGGCTCTTTCTTCCGACTCCGGG 60

QY 71 AGCGAGGCTCCCAAGTCGGAGGCGCAAGAGATGTTCTCTCTCAAGAAGTGAAACGGCGTA 130
    |||
DB 61 AGCGAGGCTCCCAAGTCGGAGGCGCAAGAGATGTTCTCTCTCAAGAAGTGAAACGGCGTA 120

QY 131 GCCATGTGAGAGCTGGAGACCTTGAAGTCGATACCTGTGCCATCTGCAGGGTCCAGGTGATG 190
    |||
DB 121 GCCATGTGAGAGCTGGAGACCTTGAAGTCGATACCTGTGCCATCTGCAGGGTCCAGGTGATG 180

QY 191 GATGCTGCTTGGAGATGTAAGTGAAGAAACAAGCAAGAGAGAGCTGTGTGTGCTTGGGGA 250
    |||
DB 181 GATGCTGCTTGGAGATGTAAGTGAAGAAACAAGCAAGAGAGAGCTGTGTGTGCTTGGGGA 240

QY 251 GAGTGAACCATTCCTTCCCAACTGCTGATGCTCCCTGGGTGAAACGAACATCCG 310
    |||
DB 241 GAGTGAACCATTCCTTCCCAACTGCTGATGCTCCCTGGGTGAAACGAACATCCG 300

QY 311 TGCCCTCTGTGCGAGAGAGCTGGGTAGTCCAAAGATCGCAAAATGAGAGGTGGCCAG 370
    |||
DB 301 TGCCCTCTGTGCGAGAGAGCTGGGTAGTCCAAAGATCGCAAAATGAGAGGTGGCCAG 360

```

```

QY 371 GCGCTCCTGCTGTGGTGTGCTGACCTCGACCAAGAATAACACTGACGGGATTCCT 430
    |||
DB 361 GCGCTCCTGCTGTGGTGTGCTGACCTCGACCAAGAATAACACTGACGGGATTCCT 420

QY 431 TGAGAGAGAGAGATGCTGTGGCCTTTGAGACTGACCAAGAGCTTTTAAATTTG 490
    |||
DB 421 TGAGAGAGAGAGATGCTGTGGCCTTTGAGACTGACCAAGAGCTTTTAAATTTG 480

QY 491 TCTGTTAGTGTGGGAAATTTCTTCAATTAAATTTGTTAAATTTGCTTCT 550
    |||
DB 481 TCTGTTAGTGTGGGAAATTTCTTCAATTAAATTTGTTAAATTTGCTTCT 540

QY 551 ACCCTGGTGTGTGTGTGATACGAAATGATAGAAAGAGCGAAGACCAAGAAATGATC 610
    |||
DB 541 ACCCTGGTGTGTGTGTGATACGAAATGATAGAAAGAGCGAAGACCAAGAAATGATC 600

QY 611 TTTGTTATCTGTACCAAGACT---GGAACATTTGTTACAGAAAGAAATTTGTTG 667
    |||
DB 601 TTTGTTATCTGTACCAAGACTGTGGGAAATTTGTTTCAAGAAAGAAATTTGTTG 660

QY 668 TTTATGCTTGTAGGCTTAAATAATGATTAACAAATGTTACAGTAAATAATGCAAT 727
    |||
DB 661 TTTATGCTTGTAGGCTT---AAAAATGATTAACAAATGTTACAGTAAATAATGCAAT 719

QY 728 GAAAGCGGACCTCTCTTAATCTTTTGTGTGGAGAGAGCAAGCGACACCTG 787
    |||
DB 720 GAAAGCGGACCTCTCTTAATCTTTTGTGTGGAGAGAGCAAGCGACACCTG 779

QY 788 CTGTCTTCAATTTGCTGTGAATGAGATTTTAACTGCACTGAGTAAGAGGCTAATCT 847
    |||
DB 780 CTGTCTTCAATTTGCTGTGAATGAGATTTTAACTGCACTGAGTAAGAGGCTAATCT 839

QY 848 CCGGTAACTGTAATATGCGCTTAATCTGTGGGTAAACGCTTTGTCTCTGACTTCTCA 907
    |||
DB 840 CCGGTAACTGTAATATGCGCTTAATCTGTGGGTAAACGCTTTGTCTCTGACTTCT-CA 897

QY 908 TCTTTGACTTGGCCAGGA 925
    |||
DB 898 TCTTTGACTTGGCCAGGA 915

```

BASE COUNT

ORIGIN

Query Match 75.1%; Score 856.4; DB 14; Length 988;
Best Local Similarity 98.7%; Pred. No. 5.7e-173;
Matches 906; Conservative 0; Mismatches 6; Indels 6; Gaps 4;

QY 11 GCGCCATGGCCGACCTGGAGGACCGCGAGGAAACCTGGCTCTTTCTTCCGACTCCGGG 70

DB 1 GCGCCATGGCCGACCTGGAGGACCGCGAGGAAACCTGGCTCTTTCTTCCGACTCCGGG 60

QY 71 AGCGAGGCTCCCAAGTCGGAGGCGCAAGAGATGTTCTCTCTCAAGAAGTGAAACGGCGTA 130

DB 61 AGCGAGGCTCCCAAGTCGGAGGCGCAAGAGATGTTCTCTCTCAAGAAGTGAAACGGCGTA 120

QY 131 GCCATGTGAGAGCTGGAGACCTTGAAGTCGATACCTGTGCCATCTGCAGGGTCCAGGTGATG 190

DB 121 GCCATGTGAGAGCTGGAGACCTTGAAGTCGATACCTGTGCCATCTGCAGGGTCCAGGTGATG 180

QY 191 GATGCTGCTTGGAGATGTAAGTGAAGAAACAAGCAAGAGAGAGCTGTGTGTGCTTGGGGA 250

DB 181 GATGCTGCTTGGAGATGTAAGTGAAGAAACAAGCAAGAGAGAGCTGTGTGTGCTTGGGGA 240

QY 251 GAGTGAACCATTCCTTCCCAACTGCTGATGCTCCCTGGGTGAAACGAACATCCG 310

DB 241 GAGTGAACCATTCCTTCCCAACTGCTGATGCTCCCTGGGTGAAACGAACATCCG 300

QY 311 TGCCCTCTGTGCGAGAGAGCTGGGTAGTCCAAAGATCGCAAAATGAGAGGTGGCCAG 370

DB 301 TGCCCTCTGTGCGAGAGAGCTGGGTAGTCCAAAGATCGCAAAATGAGAGGTGGCCAG 360

FEATURES

source

```

1..949
/organism="Mus musculus"
/mol_type="mRNA"

```

RESULT 7

LOCUS

BU840985 949 bp mRNA linear EST 16-OCT-2002
AGENCOURT_10187674 NIH_MGC_134 Mus musculus cDNA clone
IMAGE:518791 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L14M14101 row: b column: 08
High quality sequence stop: 658.

FEATURES

source

```

1..949
/organism="Mus musculus"
/mol_type="mRNA"

```

```

/db_xref="taxon:10090"
/clone="IMAGE:6518791"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 134"
/notes="Vector: pCMV-Sport6.1.ccd; Site 1: EcoRV; Site 2:
NotI; Cloned unidirectionally. Primer: Oligo dt. Average
insert size 1.7 kb. Constructed by Resgen, Invitrogen
Corp. Note: this is a NIH MGC Library."
BASE COUNT      255 a      187 c      246 g      259 t      2 others
ORIGIN

Query Match      73.4%; Score 836.2; DB 13; Length 949;
Best Local Similarity 96.2%; Pred. No. 1.2e-168;
Matches 910; Conservative 0; Mismatches 29; Indels 7; Gaps 5;

QY 82 CAAGTCGGAGGCGCAAGATGTTCTCTCAAGAAAGTGAACGGGTAGCAGTATGAGG 141
   |||||
DB 4 CAAGTCGGAGGCGCAAGATGTTCTCTCAAGAAAGTGAACGGGTAGCAGTATGAGG 63
   |||||

QY 142 CTGGGACGTTGAGTGCGATACCTGTCATCTGACAGGCTCAAGTATGATGCTGCT 201
   |||||
DB 64 CTGGGACGTTGAGTGCGATACCTGTCATCTGACAGGCTCAAGTATGATGCTGCT 123
   |||||

QY 202 TCGATGTCAGCTGAAAAACAAGCAAGAGACCTGTGTGTGTGTGGGAGAGTGAACCA 261
   |||||
DB 124 TCGATGTCAGCTGAAAAACAAGCAAGAGACCTGTGTGTGTGTGGGAGAGTGAACCA 183
   |||||

QY 262 TTCCTTCCACACCTGTCATGTCCTGCGGGTGAACCAATCGCTGCTCTGTG 321
   |||||
DB 184 TTCCTTCCACACCTGTCATGTCCTGCGGGTGAACCAATCGCTGCTCTGTG 243
   |||||

QY 322 CCAGCAGAGCTGGTGTAGTCCAAAGATCGGCAATGAGAGTGTGCCAGGCTCTGTGT 381
   |||||
DB 244 CCAGCAGAGCTGGTGTAGTCCAAAGATCGGCAATGAGAGTGTGCCAGGCTCTGTGT 303
   |||||

QY 382 GTGTGTGTCGACCTGTCGACCAAGACTAACTGACGAGGAGTTCATCTTGAAGAGAGA 441
   |||||
DB 304 GTGTGTGTCGACCTGTCGACCAAGACTAACTGACGAGGAGTTCATCTTGAAGAGAGA 363
   |||||

QY 442 GGATGCTGTGCGCTTGTGAGACTCAACAAGGCTTCTTAATTTGCTGTGTAAT 501
   |||||
DB 364 GGATGCTGTGCGCTTGTGAGACTCAACAAGGCTTCTTAATTTGCTGTGTAAT 423
   |||||

QY 502 TTGGGAAATTTCTTACAAATTAAGATAATTTGTTAAATGCGCTTCTTACCTGTGTGT 561
   |||||
DB 424 TTGGGAAATTTCTTACAAATTAAGATAATTTGTTAAATGCGCTTCTTACCTGTGTGT 483
   |||||

QY 562 GTGTGTGTCGATGCAATGCAATGAAGAGAGAACCCAGAAATGATCTTTGTTATCT 621
   |||||
DB 484 GTGTGTGTCGATGCAATGCAATGAAGAGAGAACCCAGAAATGATCTTTGTTATCT 543
   |||||

QY 622 GTACCCAGCACT--GGAACATTTGTTCACAGAAAGAAATGTTGTTATGCTTGA 678
   |||||
DB 544 GTACCCAGCACTGTGGGAACATTTGTTCACAGAAAGAAATGTTGTTATGCTTGA 603
   |||||

QY 679 GGGTTAAAAAATGATTAACGAATGTTACGATTAACAATTAATGATGAAAAGCCGAC 738
   |||||
DB 604 GGGTTAAAAAATGATTAACGAATGTTACGATTAACAATTAATGATGAAAAGCCGAC 663
   |||||

QY 739 TCTCTCTAATCC--TTTTGTTGGGAGAGAGCGACGCGCACTGCTGTCTTCAAT 797
   |||||
DB 664 TCTCTCTAATCC--TTTTGTTGGGAGAGAGCGACGCGCACTGCTGTCTTCAAT 723
   |||||

QY 798 TTGCTGTGATGAGATTTTAACTGCACTCACTGAGAGAGCGTAACTGTGGGTAAC 857
   |||||
DB 724 TTGCTGTGAGAAATGAAATTTTAACTGCACTCACTGAGAGAGCGTAACTGTGGGTAAC 783
   |||||

QY 858 GTTAATATGGGTAACCTGTGGG--AAAGCGCTTGTCTCTGACTTCTCATCTTTGACT 916
   |||||
DB 784 GTTAATATGGGTAACCTGTGGG--AAAGCGCTTGTCTCTGACTTCTCATCTTTGACT 843
   |||||

QY 917 TTGCAAGG--AAGCGTGAATGTTCAACCACTTAGTTCTAAAGAACTGTTTCTGTTTGG 975
   |||||

```

```

DB 844 TTGGCAGAAAACCTGGAATGTTCAACCACTGTTTAAAGAACTTTTCCGTTTG 903
   |||||
QY 976 CCGAAGGTTGATGTTAT--GTTTATGCAAAAATATTATGAGGAAA 1020
   |||||
DB 904 CCAAGGTTGATGATGATGTTTATGAGCCCAAAATATTAGTTAGAAA 949
   |||||

RESULT 8
CB204521
LOCUS
DEFINITION
AGENCOURT 11276019 NIH MGC 135 Mus musculus cDNA clone
IMAGE:30138707 5', mRNA sequence.
CB204521
CB204521.1 GI:28242067
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 901)
NIH-MGC http://mgi.nci.nih.gov/
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: NDAM0041 row: p column: 12
High quality sequence stop: 678.
Location/Qualifiers
1..901
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30138707"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 135"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5,
and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dT. cDNA
enrichment: >1k bp, Average insert size 1.6k bp.
Normalization (Cot value): 7.5 kb. Priming sequence:
5'GACTAGTTTATGATGCGCAGCGCGGCCGCT(3' Tissue contributed by
David Rowe. Library constructed by Resgen, Invitrogen
Corp."
BASE COUNT      231 a      190 c      241 g      239 t
ORIGIN

Query Match      73.1%; Score 833.6; DB 14; Length 901;
Best Local Similarity 98.3%; Pred. No. 4.4e-168;
Matches 885; Conservative 0; Mismatches 9; Indels 6; Gaps 4;

QY 21 CCGAGCTGAGAGAGCGGAGAGAACCTGCTCTTCTTCCACTCCGAGAGCGAGGCT 80
   |||||
DB 1 CCGAGCTGAGAGAGCGGAGAGAACCTGCTCTTCTTCCACTCCGAGAGCGAGGCT 60
   |||||

QY 81 CCAAGTCGGAGAGCGACAAGATGTTCTCTCAAGAAATGGAACCGGTTAGCCATGTGA 140
   |||||
DB 61 CCAAGTCGGAGAGCGACAAGATGTTCTCTCAAGAAATGGAACCGGTTAGCCATGTGA 120
   |||||

QY 141 GCTGGGAGGTTGAGAGGATACCTGTGCACTCTGAGAGGTCAAGGTGATGATGCTGCC 200
   |||||
DB 121 GCTGGGAGGTTGAGAGGATACCTGTGCACTCTGAGAGGTCAAGGTGATGATGCTGCC 180
   |||||

QY 201 TTGATGTCAGGCTGAAAAACAAGCAAGAGACCTGTGTGTGTGTGGGAGAGTGAAC 260
   |||||

```

Db	1b1	TTCCATGTCACGCTGAAAAACAAGACAGAGACTGTGTGTGTGTCTGTGGGAGAGGTAAAC	240
QY	261	ATTCTTCCACACATGCTGCATGTCCCTGTGGGTGAAAACAGAACATCGCTGCCTGT	320
Db	241	ATTCTTCCACACATGCTGCATGTCCCTGTGGGTGAAAACAGAACATCGCTGCCTGT	300
QY	321	GCCAGACAGACTGGGTACTCCAAAGAAATTCGGAAATGAGAGGTGGCCAGAGCGTCTCG	380
Db	301	GCCAGACAGACTGGGTACTCCAAAGAAATTCGGAAATGAGAGGTGGCCAGAGCGTCTCG	360
QY	381	TGTGGTCTGACCCCTGGACAAAGACTTAAACCTCAGGGGATTGATCCTTGAGAGAG	440
Db	361	TGTGGTCTGACCCCTGGACAAAGACTTAAACCTCAGGGGATTGATCCTTGAGAGAG	420
QY	441	AGGATGCTGTGGCGCCTTTGAGACTCACCMAAGCCTTGCTTAATTAATTTGTCGTGTAGT	500
Db	421	AGGATGCTGTGGCGCCTTTGAGACTCACCMAAGCCTTGCTTAATTAATTTGTCGTGTAGT	480
QY	501	TTTGGGAATTTCTCTACAAATTAAGTAATTTGTCTTAAATATGCGCTTTCTTACTCTGGT	560
Db	481	TTTGGGAATTTCTCTACAAATTAAGTAATTTGTCTTAAATATGCGCTTTCTTACTCTGGT	540
QY	561	TGTGTGTGTGATACGAATGACATAGAAAGCGAGAACACACAGAAATGATCTTGTATTC	620
Db	541	TGTGTGTGTGATACGAATGACATAGAAAGCGAGAACACACAGAAATGATCTTGTATTC	600
QY	621	TGTACCCACGACT--GGACATTTGTGTTCACAGAAACAATTTGTGTATGCTTG	677
Db	601	TGTACCCACGACTGTGGGACACTTGTGTTCACAGAAACAATTTGTGTATGCTTG	660
QY	678	AGGTTAAATTAATGATATAACGATGTTCACGTAAACAATTAATGATGAAAGCCGA	737
Db	661	AGGTTAAATTAATGATATAACGATGTTCACGTAAACAATTAATGATGAAAGCCGA	720
QY	738	CTCCGCCCT--AATCCCTTTTGTGTGGGAGAGAGGAACCGAGGCC--ACCGCTGTCTTC	795
Db	721	CTCCGCCCTAATCCCTTTTGTGTGGGAGAGAGGAACCGAGGCCACCGCTGTCTTC	780
QY	796	ATTGCTGTGATGAGATTTTAAACCTGCACCTCACTGAAAGAGCGCTAATCTGCGGTAA	855
Db	781	ATTGCTGTGATGAGATTTTAAACCTGCACCTCACTGAAAGAGCGCTAATCTGCGGTAA	840
QY	856	CTGTAAATAGCGGTACTGT--CGGGTAAACGGCTTTGTCTCTCGACTTCTCCATCTTGA	914
Db	841	CTGTAAATAGCGGTACTGTCCGGTAAACGGCTTTGTCTCTCGACTTCTCCATCTTGA	900

RESULT 9

LOCUS B0964159

DEFINITION AGNCOURT 10051203 NIH MGC 134 Mus musculus cDNA clone

ACCESSION B0964159

VERSION B0964159.1

KEYWORDS GI:22379637

SOURCE EST.

ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1185)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished

JOURNAL Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov

COMMENT Tissue Procurement: Dr. David Rowe cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

1185 bp mRNA linear EST 21-AUG-2002

Plate: LLAM14077		row: h	column: 07
High quality sequence stop: 704.			
Location/Qualifiers			
1. .1185			
/organism="Mus musculus"			
/mol_type="mRNA"			
/db_xref="taxon:10090"			
/clone="IMAGE:6509718"			
/cisue_type="undifferentiated 11nb"			
/lab_host="DH10B (phage-resistant)"			
/clone_11b="NIH MGC 134"			
/note="Vector: pCMV-SPORT6.1.cdb; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH MGC Library."			
BASE COUNT			
286 a 256 c 331 g 307 t 5 others			
ORIGIN			
Query Match 70.4%; Score 802.4; DB 13; Length 1185;			
Best Local Similarity 94.8%; Pred. No. 2.1e-161;			
Matches 929; Conservative 0; Mismatches 33; Indels 18; Gaps 9			
Qy	1	GTTCTGCGCGCGCCATGGCCGAGTGTGAGAGACGGACGAAACCTGCGCTTCTTC	60
Db	66	GTTCTGCGCGCGCCATGGCCGAGTGTGAGAGACGGACGAAACCTGCGCTTCTTC	1233
Qy	61	GCACTCCGGAGCGCAGGCTCCAAAGTCGGAGCGACAAAGTCTCTCTCAAGAAATG	1202
Db	126	GCACTCCGGAGCGCAGGCTCCAAAGTCGGAGCGACAAAGTCTCTCTCAAGAAATG	1855
Qy	121	GAACGCGGTAGCCATGTGAGTGGAGCTTGAGCTGAGTGCATTCCTGCACTGACGGT	1808
Db	186	GAACGCGGTAGCCATGTGAGTGGAGCTTGAGCTGAGTGCATTCCTGCACTGACGGT	2453
Qy	181	CCAGGTGATGATGCTGCTGCATGTCCATGTCAAGCTGAAACCAAGAAAGACGTGTGT	2406
Db	246	CCAGGTGATGATGCTGCTGCATGTCCATGTCAAGCTGAAACCAAGAAAGACGTGTGT	3058
Qy	241	GGTCTGGGAGAGTGTAAACCATTCCTTCACAACTGCTGATGCTCCGTGGGTGAACA	3000
Db	306	GGTCTGGGAGAGTGTAAACCATTCCTTCACAACTGCTGATGCTCCGTGGGTGAACA	3653
Qy	301	GAACAATCCGTGCTCTGTGTCACAGACGACTGGGATGCCAAGAAATCGGCAATGAGA	3606
Db	366	GAACAATCCGTGCTCTGTGTCACAGACGACTGGGATGCCAAGAAATCGGCAATGAGA	4259
Qy	361	GGTGGCCACAGGCGCTCTGTGTGTGGTGTCTGACCTCGAACAAACATCTGCAGGG	4220
Db	426	GGTGGCCACAGGCGCTCTGTGTGTGGTGTCTGACCTCGAACAAACATCTGCAGGG	4853
Qy	421	GATTCACTCTTGAGAGAGAGAGATGCTGTGGGCTTTGAGACTCATCAAAAGCTTCTT	4806
Db	486	GATTCACTCTTGAGAGAGAGAGATGCTGTGGGCTTTGAGACTCATCAAAAGCTTCTT	5459
Qy	481	TATTAATTTGTCTGTTAGTTTGGGAAATTTCTTCAATTAAGATTAATTTGTTAAAT	5406
Db	546	TATTAATTTGTCTGTTAGTTTGGGAAATTTCTTCAATTAAGATTAATTTGTTAAAT	6059
Qy	541	GCGCTTCTTACTCTGTGTGTGTGTGTGATGAATGCTATGAAGCGGAACCA	6000
Db	606	GCGCTTCTTACTCTGTGTGTGTGTGTGATGAATGCTATGAAGCGGAACCA	6653
Qy	601	GAATAATGATCTTGTGTTATCTGTACCCAGCACT---GGAACATGTGTTACAGAAAGAC	6577
Db	666	GAATAATGATCTTGTGTTATCTGTGTACCCAGCACTGTGGGACATGTGTTCACAGAGAC	7233
Qy	658	ATTGTTGTGTTATGCTTGAAGGTTAAAAATAGATTAAGAAATGTTACGTAACAAT	7177
Db	726	ATTGTTGTGTTATGCTTGAAGGTTAAAAATAGATTAAGAAATGTTACGTAACAAT	7853
Qy	718	AAAATGCAATTAAGAAAGCGACTCTCTCTTAATCC--TTTGTGTGTTGGAGAGAGGCAAGC	7755
Db	786	AAAATGCAATTAAGAAAGCGACTCTCTCTTAATCCNNTTTGTGTGTTGGAGAGAGGCAAGC	8455

OY	776	AAGG-CAACCCCTGCTGCTTTCATTTTGGCTGTG--AATAGAGATTTTAACTTGACACTGAGT	832
Db	846	GAGGCCCAACCTGTGCGANGCATATTTGTGGGGAATGANGGATTTTAACTTGACACTGAGT	905
OY	833	AAG-AGGGCTACTGTGCGGTAAACTG---TAATATGCGCTGACTGTG---GGGTTAAACG	885
Db	906	AAGAAAGCGTACTGTGCGGGTAACTGGAAAAATATGGGCCGTAACTGTGCGGGCTTAAACGG	965
OY	886	GCTTTGTCTCTGTG-ACTTCTCCATCTTTGACTTGGCCA--GGAAGCCGTGATTGTCAAC	942
Db	966	CTTTGTCTCTGAACTCTTCATCTTTGACTTGGCCAAAGAAAACTGGATTGTGTTCAAC	1025
OY	943	CACCTAGCTTTAAAGAACTG	962
Db	1026	CACCTAGCTTTTAAAGAACAG	1045

RESULT 10						
B0964115						
LOCUS	B0964115	968 bp	mRNA	linear	EST 21-AUG-2002	
DEFINITION	AGENCOURT_10053663 NIH MGC 134 Mus musculus					
	IMAGE:6514192 5',					
	mRNA sequence.					

ACCESSION	BQ964115	GI:22379593
VERSION	BQ964115.1	
KEYWORDS	EST	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

REFERENCE
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia, Eutheria, Rodentia; Sciurognathi, Muridae, Mus-
1 (bases 1 to 968)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC).
Unpublished
Contact: Robert Strusberg, Ph.D.

Email: cgaaps-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://Image.lnl.gov>
Plate: LLM14089 row: b column: 17
High quality sequence stop: 666.

```

FEATURES
  high quality sequence scop: 000.
  location/Qualifiers
    1. .968
source

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6514192"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH_MGC_134"
/notes="Vector: pCMV-SPORT6.1.ccdB. Site_1: EcoRV, Site_2:
NotI; Cloned unidirectionally. Primer: Oligo dT. Average
insert size 1.7 kb. Constructed by ResGen, Invitrogen
Corp. Note: this is a NH MGC Library."
BASE COUNT      288 a      170 c      219 g      282 t      9 others

```

Query Match	68.6%	Score 781.8;	DB 13;	Length 968;
Best Local Similarity	98.8%	Pred. No. 5,5e-157;		
Matches 840;	Conservative	0;	Mismatches 3;	Indels 7;
			Gaps	5;

QY 293 GTGAAACGAGACATCGCTGCTCTGTGGCAGACGACTGGGTAGTCCAAAGATCGGC 352

Db 1 GTGAAACGAGACATCGCTGCTCTGTGGCAGACGACTGGGTAGTCCAAAGATCGGC 60

QY 353 AAATGAGAGGTGGCCACGAGCGCTCTCTGTTGTTGCTGACCCCTGGACAAAGACTAAACA 412

Db 61 AAATGAGAGGTGGCCACGAGCGCTCTCTGTTGTTGCTGACCCCTGGACAAAGACTAAACA 120

Oy	413	CTGAGGGGATTCATCTCTTAGAGAGAGAGATGCTGTGGCCCTTTGAGACTACCCAAAG	472
Db	121	CTGAGGGGAGATCATCTCTTAGAGAGAGAGAGATGCTGTGGCCCTTTGAGACTACCCAAAG	180
Oy	473	GCTTGCTTTATTAATTTGTCTGTTAAGTTTGGGAAATTCCTACAAATTAGATTAATTTG	532
Db	181	GCTTGCTTTATTAATTTGTCTGTTAAGTTTGGGAAATTCCTACAAATTAGATTAATTTG	240

QY 533 TTTAAATGGCCTTCTACCCCTGTGTGTGTGATACGAATCATACGAACGCGA 592
Db 241 TTTAAATGGCCTTCTCTACCTGTGTGTGTGTGTGATACGAATCATACGAACGCGA 300

301 GAACACCAAGAAATGATCTTGTATTACTGTACCCACGACACTGTGGCAATGTGTTCAC 360
 353 GAACACCCAGAAATGATCTTGTATTACTGTACCCACGACACTGTGGCAATGTGTTCAC 649
 359 GAACACCCAGAAATGATCTTGTATTACTGTACCCACGACACTGTGGCAATGTGTTCAC 655

Db 361 AGAAGACATTGTTGTTGTTATGCTTGAGGGTTAAAAATGATAACGAATGTTACAG 420

421 TACCAATTAATGCAATGAAAGCGACTCTCTTATCTTTTGTGTGGAGAGAG 480

Db	481	GCAGCCAGGCCACCCCTGCTCTTCAATTGCTGATATAGGATTTTAACTGCATCTCA	540
Oy	830	GTTAAGAGGCGTACTGTCTGGGGTAACTGTATATATGCGCGTAACTGTCTGGGTAAACGGCTT	889

Dd 541 GTGAAGAGCGCTAACTGTCGGGTAAACTGTATATATGCGCTAACTGTCGGGTAAACGGCTT 600

Dy 890 TGTCTCCTGCATTCTCCATCTTTTGACTTGGCCAGAGAACCTGCATTTCTCAACCATTTAG 949

DB 601 TGGCTCCCTGACCTCTCCATCTTTGACTTGGCCAGGAAAGCCCTGGATTGTTCAACCAACTTAC 660

OY 950 TTCTAAGAACTGTTCTCTGTTTTGGCGAAGTGTGATTTGATGTTTTAGTCGCAAAATA 1009

1010 TTAGAGGAAATGGCTTACTAGTATPAACACTGAGCTTCATTATGCAATGTTTAAATPAA 1069

Oy	1070	AAATTG- <u>GCTTGAGTTAATAAGTTGATATACTCTTAAATCATTAACGTAATTC</u>	1128
Db	778	ATTTCGCGCTTTGACCTTTTATAGCTGGATTTCTCTCTTTTAACTTTTTCCTTTTC	827

QY	1129	ATCAATTAA	1138
Db	838	ATCAATTAA	847

RESULT 11
B1732003

DEFINITION	603355869F1.NH.MGC_94 Mus musculus cDNA clone IMAGE:5363056 5', mRNA sequence.
ACCESSION	B1732003
VERSION	B1732003.1 GI:15709016

GENUS	SOURCE	ORGANISM
ES1	Mus musculus (house mouse)	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mus musculus	Mus musculus	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mus musculus	Mus musculus	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 820)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT:
Contact: Robert Strausberg, Ph.D.
Email: cgsab@r-mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.

Db 67 GAGGCGAAGATGTTCCCTCTCTCAAGAAAGTGAACGGGATGACCATGTGAGCGTGGAC 126
 Qy 149 GTTGAAGTGAATACCTGTTGATCCATGCGAGGGTCCAGTGTATGATGCTGCTTCATGT 208
 Db 127 GTTGAAGTGAATACCTGTTGATCCATGCGAGGGTCCAGTGTATGATGCTGCTTCATGT 186
 Qy 209 CAAGCTGAAAACAAGCAAG 268
 Db 187 CAAGCTGAAAACAAGCAAG 246
 Qy 269 CACACTGTGTATGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 328
 Db 247 CACAACTGTGTATGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 306
 Qy 329 GACTGGGTGATCCAAAGATCGGCAAAATGAGAGGTGGCCAGGCGCTCTGTGTGTGTTG 388
 Db 307 GACTGGGTGATCCAAAGATCGGCAAAATGAGAGGTGGCCAGGCGCTCTGTGTGTGTTG 366
 Qy 389 CTGACCTGTGACAAAGACTAAACACTGACGGGGATTCATCTTTGAGAGAGAGAGATGCT 448
 Db 367 CTGACCTGTGACAAAGACTAAACACTGACGGGGATTCATCTTTGAGAGAGAGATGCT 426
 Qy 449 GGGCGCTTTGAG 508
 Db 427 GGGCGCTTTGAG 486
 Qy 509 ATTCTCTCAATTAAGATATTTGTTAAATAATGCGCTTTCCTACCTGCTGTGTGTGTG 568
 Db 487 ATTCTCTCAATTAAGATATTTGTTAAATAATGCGCTTTCCTACCTGCTGTGTGTGTG 546
 Qy 569 TGAATACGATGATGAG 628
 Db 547 TGAATACGATGATGAG 606
 Qy 629 CGACT---GGAACATGTTGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685
 Db 607 CGACTGTTGGAACATGTTGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
 Qy 686 AAAATGATTAAGATGTTTAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 743
 Db 667 AAAATGATTAAGATGTTTAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 726
 Qy 744 CTATATCTTTTGT 802
 Db 727 CTATATCTTTTGT 786
 Qy 803 G--TGAATGAGATTTAACT--GCACTGAGTGAAG--AGGCGTAACTGCGGTTAACTG 858
 Db 787 GCGGAGATGAGATTTAACTGCGGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
 Qy 859 TAATATGCGGTAACTGTGCGGTTAAACGCGCTTGTCTC 895
 Db 847 TAATATGCGGTAACTGTGCGGTTAAACGCGCTTGTCTC 883
 RESULT 13
 CA980368 882 bp mRNA linear EST 06-JAN-2003
 LOCUS CA980368
 DEFINITION AGENCOURT 11294967 NIH MGC 164 Mus musculus cDNA clone
 IMAGE:30145601 5', mRNA sequence.
 ACCESSION CA980368
 VERSION CA980368.1 GI:27513022
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 882)
 AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>,
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. David Rowe and Dr. Mina
 cDNA Library Preparation: Invitrogen Corp
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: NDAM0059 row: 0 column: 18
 High quality sequence stop: 662.
 Location/Qualifiers
 1..882
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30145601"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 164"
 /note="Vector: pCMV-SPORT6.1.cdb; Site 1: EcoRV; Site 2:
 NotI; Non-normalized full-length enriched library from
 pooled mouse embryonic limb, maxilla and mandible, day
 10.5 and 11.5 (size selected for the 0.5-1 kb fragments)
 Cloned directionally, priming method: Oligo-dT. cDNA
 enrichment: >1k bp. Average insert size 1.8k bp. Priming
 sequence: 5'-GACTGTTCTGATGCGAGCGGCCCTT 3'. Tissue
 contributed by, David Rowe. Library constructed by ResGen,
 Invitrogen Corp."

BASE COUNT 234 a 173 c 230 g 244 t 1 others
 ORIGIN
 Query Match 64.9%; Score 740; DB 14; Length 882;
 Best Local Similarity 97.4%; Pred. No. 4.9e-148;
 Matches 785; Conservative 0; Mismatches 15; Indels 6; Gaps 3;

Qy 191 GATGCTGCTTGCATGCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250
 Db 30 GATGCTGCTTGCATGCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 89
 Qy 251 GAGGTAACTTCCTCCCACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 310
 Db 90 GAGGTAACTTCCTCCCACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 149
 Qy 311 TGCCCTGTGTCAG 370
 Db 150 TGCCCTGTGTCAG 209
 Qy 371 GCGCTCTGT 430
 Db 210 GCGCTCTGT 269
 Qy 431 TGAG 490
 Db 270 TGAG 329
 Qy 491 TCTGTTTATGTTTGGGAATTTCTACATTAATTAATTAATTAATTAATTAATTAATTAAT 550
 Db 330 TCTGTTTATGTTTGGGAATTTCTACATTAATTAATTAATTAATTAATTAATTAATTAAT 389
 Qy 551 ACCCTGT 610
 Db 390 ACCCTGT 449
 Qy 611 TTTGTTTATCTGTACCCAGACT---GGAACATGTTGTTCAAGAGAGAGAGAGAGAGAGAG 667
 Db 450 TTTGTTTATCTGTACCCAGACTGTGGGAACATGTTGTTCAAGAGAGAGAGAGAGAGAGAG 509
 Qy 668 TTTATGCTTGAAGGTTAAAAATAGATTAACGAATTTTACATTAACGAATTTTATTAATGCT 727
 Db 510 TTTATGCTTGAAGGTTAAAAATAGATTAACGAATTTTACATTAACGAATTTTATTAATGCT 569
 Qy 728 GAAAAGCCGACTCTCTCTTAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 787
 Db 570 GAAAAGCCGACTCTCTCTTAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 629

QY 788 CTGCTTCATTTCTGTAATGAGATTTTAACCTGACCTGATGAAGAGCGTAACTGT 847
 |||||
 Db 630 CTCTCTTCTTCTGCTGTAATGAGATTTTAACCTGACCTGATGAAGAGCGTAACTGT 689
 |||||
 QY 848 CCGGTAACCTGTAATATGCGTAACTGTGCGGTAAACGGCTTTGTCTCTGACTTCTCA 907
 |||||
 Db 690 CCGGTAACTGTAATATGCGTAACTGTGCGGTAAACGGCTTTGTCTCTGACTTCTCA 749
 |||||
 QY 908 TCTTTGACTTGGCCAGGAA-GCCTGATTTGTTCAACCACTTAGTT--CTAAGAAGTGT 964
 |||||
 Db 750 TCTTTGACTTGGCCAGGAAAGCGCGGATTTGTTCCACCACTTAGTTCTAAGAAGCGGTT 809
 |||||
 QY 965 TTCTGTTTGGCCAGGATTTGTTATG 990
 |||||
 Db 810 TTCCGGTTTGGCCAAAGGTGGAATTTG 835
 |||||

RESULT 14
 CB209388 938 bp mRNA linear EST 05-FEB-2003
 LOCUS CB209388
 DEFINITION IMAGE:30245249 5', mRNA sequence.
 ACCESSION CB209388
 VERSION CB209388.1 GI:28250951
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://imgc.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. David Rowe and Dr. Mina
 CDNA Library Preparation: Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: NDAM0319 row: 9 column: 18
 High quality sequence stop: 712.
 Location/Qualifiers
 1..938
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30245249"
 /lab_host="PH10B (phage-resistant)"
 /note="Vector: PCMV-SPORT.1.cdb; site 1: EcoRV; site 2:
 NotI; Non-normalized full-length enriched library from
 pooled mouse embryonic limb, maxilla and mandible, day
 10.5 and 11.5 (size selected for the 0.5-1 kb fragments)
 Cloned directionally, priming method: Oligo-dT. CDNA
 enrichment: >1k bp. Average insert size 1.8k bp. Priming
 sequence: 5'-GACTAGTTCTAGATCGGAGCGGCCCTT 3'. Tissue
 contributed by: David Rowe. Library constructed by Resgen,
 Invitrogen Corp."

BASE COUNT 204 a 153 c 211 g 202 t 168 others
 ORIGIN
 Query Match 64.9%; Score 739.6; DB 14; Length 938;
 Best Local Similarity 99.1%; Pred. No. 66-148;
 Matches 755; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 67 CCGGAGCGGAGGCTCCAGTCCGGAGCGGCAAGATGTTCTCTCAAGAGTGAAGCC 126
 |||||
 Db 1 CCGGAGCGGAGGCTCCAGTCCGGAGCGGCAAGATGTTCTCTCAAGAGTGAAGCC 60
 |||||

QY 127 GGTAGCCATGTGAGAGCTGGAGCGTTGAGTGCGATACCTTGCCATCTGAGGGTCCAGGT 186
 |||||
 Db 61 GGTAGCCATGTGAGAGCTGGAGCGTTGAGTGCGATACCTTGCCATCTGAGGGTCCAGGT 120
 |||||
 QY 187 GATGAGTGCCTGCTTCCGATGTCGAAGCTGTAAGCAAGCAAGAGAGAGCTGTGTGTGCTG 246
 |||||
 Db 121 GATGAGTGCCTGCTTCCGATGTCGAAGCTGTAAGCAAGCAAGAGAGAGCTGTGTGTGCTG 180
 |||||
 QY 247 GGGAGAGTGAACCATTCCTTCAACAAGCTGCTGCAATGCTTGGGTGTAACAGAACAA 306
 |||||
 Db 181 GGGAGAGTGAACCATTCCTTCAACAAGCTGCTGCAATGCTTGGGTGTAACAGAACAA 240
 |||||
 QY 307 TCGCTGCCCTCTGTGTCAGAGAGAGCTGGGTAGTCCAAAGAAATCGGCAATGAGAGTGGC 366
 |||||
 Db 241 TCGCTGCCCTCTGTGTCAGAGAGAGCTGGGTAGTCCAAAGAAATCGGCAATGAGAGTGGC 300
 |||||
 QY 367 CCAGGCGCTCTGTGTGTGCTGAGACCTGTGACAAAGACTAAGACTGACAGGGATTTCA 426
 |||||
 Db 301 CCAGGCGCTCTGTGTGTGCTGAGACCTGTGACAAAGACTAAGACTGACAGGGATTTCA 360
 |||||
 QY 427 TCGTTGAGAGAGAGAGATGCTGTGCGCTTTGAGACTCAACCAAGGCTGCTTTATTTA 486
 |||||
 Db 361 TCGTTGAGAGAGAGAGATGCTGTGCGCTTTGAGACTCAACCAAGGCTGCTTTATTTA 420
 |||||
 QY 487 TTTGTCTGTTTGTGTTGGGAAATCTCTACAAATTAAGTAATTTGTTAAATGGCCCTT 546
 |||||
 Db 421 TTTGTCTGTTTGTGTTGGGAAATCTCTACAAATTAAGTAATTTGTTAAATGGCCCTT 480
 |||||
 QY 547 TCTTACCTCTGT 606
 |||||
 Db 481 TCTTACCTCTGT 540
 |||||
 QY 607 GATCTTTGTTATCTGTATCTGTATCCAGCACT--GGAACATGTGTCTCACAGAAACATTGTT 663
 |||||
 Db 541 GATCTTTGTTATCTGTATCTGTATCCAGCACTGTGGAAACATTGTGTCTACAGAAACATTGTT 600
 |||||
 QY 664 TGTGTTTATGCTTGAAGGTTAAATAATGATTAAGAAATGTATCAAGTAACAATAAAG 723
 |||||
 Db 601 TGTGTTTATGCTTGAAGGTTAAATAATGATTAAGAAATGTATCAAGTAACAATAAAG 660
 |||||
 QY 724 CATTGAAAAGCCGACTCCTCTTAATCCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 783
 |||||
 Db 661 CATTGAAAAGCCGACTCCTCTTAATCCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
 |||||
 QY 784 CCGCTGCTGCTTATTTGCTGTGATGAGAGATTTTAACCTGCA 825
 |||||
 Db 721 CCGCTGCTGCTTATTTGCTGTGATGAGAGATTTTAACCTGCA 762
 |||||

RESULT 15
 BY710626 965 bp mRNA linear EST 16-DEC-2002
 LOCUS BY710626
 DEFINITION BY710626 RIKEN full-length enriched, 10 days embryo Mus musculus
 CDNA clone 2610005N22 5', mRNA sequence.
 ACCESSION BY710626
 VERSION BY710626.1 GI:27121853
 KEYWORDS EST.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://imgc.ncl.nih.gov/
 Tissue Procurement: Dr. David Rowe and Dr. Mina
 CDNA Library Preparation: Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: NDAM0319 row: 9 column: 18
 High quality sequence stop: 712.
 Location/Qualifiers
 1..965
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30245249"
 /lab_host="PH10B (phage-resistant)"
 /note="Vector: PCMV-SPORT.1.cdb; site 1: EcoRV; site 2:
 NotI; Non-normalized full-length enriched library from
 pooled mouse embryonic limb, maxilla and mandible, day
 10.5 and 11.5 (size selected for the 0.5-1 kb fragments)
 Cloned directionally, priming method: Oligo-dT. CDNA
 enrichment: >1k bp. Average insert size 1.8k bp. Priming
 sequence: 5'-GACTAGTTCTAGATCGGAGCGGCCCTT 3'. Tissue
 contributed by: David Rowe. Library constructed by Resgen,
 Invitrogen Corp."

BASE COUNT 204 a 153 c 211 g 202 t 168 others
 ORIGIN
 Query Match 64.9%; Score 739.6; DB 14; Length 938;
 Best Local Similarity 99.1%; Pred. No. 66-148;
 Matches 755; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 67 CCGGAGCGGAGGCTCCAGTCCGGAGCGGCAAGATGTTCTCTCAAGAGTGAAGCC 126
 |||||
 Db 1 CCGGAGCGGAGGCTCCAGTCCGGAGCGGCAAGATGTTCTCTCAAGAGTGAAGCC 60
 |||||

REFERENCE
 AUTHORS
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons


```

RESULT 16
CA323814
LOCUS
DEFINITION
CA323814 743 bp mRNA linear EST 26-NOV-2002
IMAGE: 6820005 5', mRNA sequence.
ACCESSION
CA323814
VERSION
CA323814.1 GI:24541912
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 743)
National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS
Unpublished
JOURNAL
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
FEATURES
source
Seq primer: pyx-5.
Location/Qualifiers
1..743
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6820005"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP FX0"
/notes="Organ: Brain; Vector: pyx-Asc; Site:1, Ecor I;
Site_2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction. Ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pyx-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the polYA tail
Iowa Brain Anatomy Project (BMAP); 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemm Chin, Ph.D.,
program coordinator."
BASE COUNT
196 a 149 c 197 g 201 t
ORIGIN
Query Match 63.6%; Score 725.4; DB 14; Length 743;
Best Local Similarity 99.5%; Pred. No. 6,6e-145;
Matches 739; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

```

```

QY 274 CTGCTCATGTCCTCTGGGTGAAAAGAAACATCGCTCCCTGTCAGACAGACTG 333
DB 181 CTGCTCATGTCCTCTGGGTGAAAAGAAACATCGCTCCCTGTCAGACAGACTG 240
QY 334 GGTAGTCCAAAGATCGCAATGAGAGGTGCCCCAGCGCTCTGCTGTTGCTGAC 393
DB 241 GGTAGTCCAAAGATCGCAATGAGAGGTGCCCCAGCGCTCTGCTGTTGCTGAC 300
QY 394 CCTGAGCAAAAGACTAAACCTGACGGGATTCATCTTGAGAGAGAGAGAGTGTGG 453
DB 301 CTTGAGCAAAAGACTAAACCTGACGGGATTCATCTTGAGAGAGAGAGAGTGTGG 360
QY 454 CCTTGAGACTACCAAAAGCTTGCTTTATTTGCTGTTAGTTTGGAAATGCT 513
DB 361 CTTTGAGACTACCAAAAGCTTGCTTTATTTGCTGTTAGTTTGGAAATGCT 420
QY 514 CTACCAATTAAATATTTGTTAAAAATGCTTCCTACCTGCTGCTGCTGTGATA 573
DB 421 CTACCAATTAAATATTTGTTAAAAATGCTTCCTACCTGCTGCTGCTGTGATA 480
QY 574 CGAATGCATAGAGAGCGGAGACACCAAGAAATGATCTTTGTTATCTGACCGACT 633
DB 481 CGAATGCATAGAGAGCGGAGACACCAAGAAATGATCTTTGTTATCTGACCGACT 540
QY 634 ---GGAACATGTTGTTCAAGAAACATTTGTTGTTATGTTAGAGGTTAAAAAT 690
DB 541 GTGGAGACATTTGTTTACAGAAAGACATTTGTTGTTATGTTAGAGGTTAAAAAT 600
QY 691 AGATAAAGCAATGTTTACAGTAACAATAAATGATGAAAGCCGACTCTCTAATCC 750
DB 601 AGATAAAGCAATGTTTACAGTAACAATAAATGATGAAAGCCGACTCTCTAATCC 660
QY 751 TTTTGTGTTGGAGAGAGGCAAGGAGCCACCTGCTGCTTCAATTTGCTGTAATA 810
DB 661 TTTTGTGTTGGAGAGAGGCAAGGAGCCACCTGCTGCTTCAATTTGCTGTAATA 720
QY 811 GGATTTTAACCTGCAGCTCAGTGA 833
DB 721 GGATTTTAACCTGCAGCTCAGTGA 743

```

```

RESULT 17
BUI13948
LOCUS
DEFINITION
AGENCOURT 10110282 NIH_MGC_134 Mus musculus cDNA clone
IMAGE:510542 5', mRNA sequence.
ACCESSION
BUI13948
VERSION
BUI13948.1 GI:22821474
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 970)
National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS
Unpublished
JOURNAL
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1M14079 row: j column: 15
High quality sequence start: 18
High quality sequence stop: 464.
FEATURES
source
1..970
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"

```

```

/db_xref="taxon:10090"
/clone="IMAGE:6510542"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 134"
/notes="Vector: PCMV-SPORE 1.1.cdb; Site 1: EcoRV; Site 2:
NotI. Cloned unidirectionally. Primer: Oligo dT. Average
insert size 1.7 kb. Constructed by Resgen, Invitrogen
Corp. Note: this is a NIH MGC library."

BASE COUNT      232 a      255 c      254 g      223 t
ORIGIN
Query Match      63.1%; Score 719.4; DB 13; Length 970;
Best Local Similarity 93.2%; Pred. No. 1.3e-143;
Matches 797; Conservative 0; Mismatches 51; Indels 7; Gaps 4;

21  CCGACGTGAGAGAGCGGCGAGAGACCTGCGTCTCTTCTTCCGACCTCCGAGAGCGAGCT 80
27  CCGACGTGAGAGAGCGGCGAGAGACCTGCGTCTCTTCTTCCGACCTCCGAGAGCGAGCT 86
81  CCAAGTCGGAGAGCGGCGAGAGAGTGTCTCTCTCAAGAGTGAACGGGTAGCCATGTGGA 140
87  CCAAGTCGGAGAGCGGCGAGAGAGTGTCTCTCTCAAGAGTGAACGGGTAGCCATGTGGA 146
141 GCTGGACGTTGAGTGGATACCTGTGCCATCTGCAGGGTCCAGGTGATGATGCTGCGC 200
147 GCTGGACGTTGAGTGGATACCTGTGCCATCTGCAGGGTCCAGGTGATGATGCTGCGC 206
201 TTGGATGTCAAGCTGAAAACACACAGAGAGAGTGTGTGGTGGGAGAGTGTAC 260
207 TTGGATGTCAAGCTGAAAACACACAGAGAGAGTGTGTGGTGGGAGAGTGTAC 266
261 ATTCTTCCACAACTGCTGCATGTCTCTGTGGTGAACAGAACATTCGCTGCTCTGT 320
267 ATTCTTCCACAACTGCTGCATGTCTCTGTGGTGAACAGAACATTCGCTGCTCTGT 326
321 GCCAGCAGAGCTGGGTGTCTCAAGAGATCGGCAATGAGAGTGGCCAGGCGCTCTCTG 380
327 GCCAGCAGAGCTGGGTGTCTCAAGAGATCGGCAATGAGAGTGGCCAGGCGCTCTCTG 386
381 TGTGTGTGTGAGACCTGTGAGCAAAAGCTAAACCTGCAGGGATTCATCTTGAAGAG 440
387 TGTGTGTGTGAGACCTGTGAGCAAAAGCTAAACCTGCAGGGATTCATCTTGAAGAG 446
441 AGGATGTGTGTGAGACCTGTGAGCAAAAGCTTCTTATTAATTTGTCTGTACT 500
447 AGGATGTGTGTGAGACCTGTGAGCAAAAGCTTCTTATTAATTTGTCTGTACT 506
501 TTTGGGAAATTTCTTACAAATTAAGATTAATTTGTTAAATATGACCTTCTACCTGTG 560
507 TTTGGGAAATTTCTTACAAATTAAGATTAATTTGTTAAATATGACCTTCTACCTGTG 566
561 TGTGTGTGTGATACGATGATAGAGAGCGAGAACACAGAGAAATGATCTTTGTTATC 620
567 TGTGTGTGTGATACGATGATAGAGAGCGAGAACACCGAGAAATGATCTTTGTTATC 626
621 TGTACCCACGACT---GGAACATTTGTCTTACAGAGAGACATTTGTTGTTATGCTG 677
627 TGTACCCCGGACTGTGGGAACTTGTGTCTCCGAGAAACAACTTTGTTGTTATGCTG 686
678 AGGATTAATAATATAGATTAAGATGTTACAGTAAACAATTAATTAATGATTAAGAGCGA 737
687 AGGATTAATAATATAGATTAAGATGTTACAGTAAACAATTAATTAATGATTAAGAGCGA 746
738 CTCCTCTAATCTCTTTTGTGTGGAGAGAGGAGGCG-AGGCCACCTTGTCTTCTTCA 796
747 CTCCTCTAATCTCTTTTGTGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 806
797 TTTGCTGTG-ATAGAGATTTTAACCTGCACTCAAGTGAAGAGGCT--AAGTGTGAGGTA 853
807 TTTGCTGTGAGAGAGATTTTAACCTGCACTCAAGTGAAGAGGAGGAGGAGGAGGAGGAG 866
854 AACTGTATATGCGC 868

```

```

Db      867 AACTGTATATGCGC 881
|||||
RESULT 18
B0571112
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: c9apds-f@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1..728
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5715295"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP FBO"
/notes="Organ: brain; Vector: pyx-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pyx-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TGAGAGAGCC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT      187 a      151 c      197 g      193 t
ORIGIN
Query Match      62.0%; Score 707.2; DB 13; Length 728;
Best Local Similarity 99.2%; Pred. No. 5.2e-141;
Matches 722; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

37  CGAGAACCTCGTCTCTTCTTCCGACCTCCGAGAGCGAGCTCCAAAGTGGAGGCGGA 96
Db   1  CGAGAACCTCGTCTCTTCTTCCGACCTCCGAGAGCGAGCTCCAAAGTGGAGGCGGA 60
97  CAAAGATGTTCTCTTCAAGAGTGAACCGGTACCAATGTGAGACTGGAGAGCTTGAGTG 156
Db   61  CAAAGATGTTCTCTTCAAGAGTGAACCGGTACCAATGTGAGACTGGAGAGCTTGAGTG 120

```

Oy		157	CGATACCTGTGCATCTGAGGGTCCAGGTATGATGCGTCCTTGATGCAAGCTGA	216
Db		121	CGATACCTGTGCATCTGAGGGTCCAGGTATGATGCGTCCTTGATGCAAGCTGA	180
Oy		217	AAACAGCAGAAGGACTGTGTGTGTGTGTGGGGAAGTGAACCATTCCTTCCAACCTG	276
Db		181	AAACAGCAGAAGGACTGTGTGTGTGTGTGGGGAAGTGAACCATTCCTTCCAACCTG	240
Oy		277	CTGCATGTCCCTGTGGGTGAAAACAGAACATGCTGCCCTCTGTGCCACAGACTGGGT	336
Db		241	CTGCATGTCCCTGTGGGTGAAAACAGAACATGCTGCCCTCTGTGCCACAGACTGGGT	300
Oy		337	AGTCCAAAGAATCGGCAATGAGAGTGGCCCCAGGCGCTCTGTGTGTGTGTGTAACCTT	396
Db		301	AGTCCAAAGAATCGGCAATGAGAGTGGCCCCAGGCGCTCTGTGTGTGTGTGTAACCTT	360
Oy		397	GCACAAAGACTAAACACTGCAGGGGATTATCTCTGAGAGAGAGATGCTGTGCGCT	456
Db		361	GCACAAAGACTAAACACTGCAGGGGATTATCTCTGAGAGAGAGATGCTGTGCGCT	420
Oy		457	TTGAGACTCACCAAAGCGTTGCTTTATTTATTTGTTCTGTGTTTAGTTTTGGAAATTCCTGA	516
Db		421	TTGAGACTCACCAAAGCGTTGCTTTATTTATTTGTTCTGTGTTTAGTTTTGGAAATTCCTGA	480
Oy		517	CAATTAAATATTTTGTAAAAATGGCCTTCTTCTCTGTGTGTGTGTGTATGCA	576
Db		481	CAATTAAATATTTTGTAAAAATGGCCTTCTTCTCTGTGTGTGTGTGTATGCA	540
Oy		577	ATGCATAGAAGCGGAGAAACCAGAAATATATCTTTTATCTGTATCCACGACT---	633
Db		541	ATGCATAGAAGCGGAGAAACCAGAAATATATCTTTTATCTGTATCCACGACTGTG	600
Oy		634	GGAACATTGTGTTCACAGAGAACATTTGTGTGTTTATGCTTGAGGGTTAAAAATAGA	693
Db		601	GGAACATTGTGTTCACAGAGAACATTTGTGTGTTTATGCTTGAGGGTTAAAAATAGA	660
Oy		694	TAAACGATGTTTACGTAACAATAATGCAATTTGATAAAAGCCGACTCTCTAATCCTTT	753
Db		661	TAAACGATGTTTACGTAACAATAATGCAATTTGATAAAAGCCGACTCTCTAATCCTTT	720
Oy		754	TTGTGTTG 761	
Db		721	TTGTGTTG 728	
RESULT 19				
AK007588				
LOCUS				
DEFINITION				
LOCATION				
FEATURES				
KEYWORDS				
SOURCE				
ORGANISM				
Mus musculus				
(house mouse)				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE				
AUTHORS				
Carninci, P. and Hayashizaki, Y.				
High-efficiency full-length cDNA cloning				
Meth. Enzymol. 303, 19-44 (1999)				
JOURNAL MEDLINE				
PUBMED				
10349636				
REFERENCE				
AUTHORS				
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,				
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
Normalization and subtraction of cap-trapper-selected cDNAs to				
prepare full-length cDNA libraries for rapid discovery of new genes				
Genome Res. 10 (10), 1617-1630 (2000)				
JOURNAL MEDLINE				
PUBMED				
11042159				
REFERENCE				
3				

AUTHORS
 Shibata, K., Itoh, M., Aizawa, K., Nagoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaiguchi, S., Ikegami, T., Tanabuchi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Kozeki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE
 RIKEN integrated sequence analysis (RISA) system-384-format
JOURNAL
 Genome Res. 10 (11), 1757-1771 (2000)
PUBMED
 11076861
REFERENCE
 4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aikawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Katsukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nkaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stambli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Bares, G., Blake, Y., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gietlinch, S., Hill, D., Hotamoni, L., Hume, D.A., Kamita, M., Lee, N.H., Lyons, P., Machionni, L., Mashima, J., Mazzarelli, J., Mommaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyok-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyshak-Boris, A., Yoshida, K., Haegawa, Y., Kawai, H., Kontecki, S. and Hayashizaki, Y.
TITLE
 Functional annotation of a full-length mouse cDNA collection
JOURNAL
 Nature 409 (6821), 685-690 (2001)
PUBMED
 12187851
REFERENCE
 5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 722)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Aikawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kakuwa, T., Kato, H., Kawai, D., Kojima, Y., Komno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Shukita, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE
 Direct Submission
JOURNAL
 Submitted (10-JUL-2000) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGGAGGACGATCCAGACCTCTTTTCTTTTNN 3']. cDNA was prepared by using crehalase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went

through one round of subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence 15', cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI. Host: SOLR.

FEATURES

source

1..722
Location/Qualifiers

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:1810023A07"
/db_xref="MGI:1892879"
/db_xref="taxon:10090"
/clone="1810023A07"
/sex="male"
/tissue_type="pancreas"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 day old"
..341
..341
/note="putative
ring finger protein 7 (MGI:1337096)"
/db_xref="MGI:1337096"
701..706
/note="putative"
722
/note="putative"

misc_feature

..341
..341
/note="putative
ring finger protein 7 (MGI:1337096)"
/db_xref="MGI:1337096"
701..706
/note="putative"
722
/note="putative"

polyA_signal

701..706
/note="putative"
722
/note="putative"

polyA_site

701..706
/note="putative"
722
/note="putative"

BASE COUNT 191 a 147 c 201 g 183 t
ORIGIN

Query Match 61.5%; Score 700.8; DB 11; Length 722;
Best Local Similarity 99.3%; Pred. No. 1.2e-139;
Matches 715; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

19 GGGCGAGGTGGAGAGCGGCGAGGAACTGCGCTCTTCTTCCGCACTCCGCGAGCGCAGG 78
2 GGGCGAGGTGGAGAGCGGCGAGGAACTGCGCTCTTCTTCCGCACTCCGCGAGCGCAGG 61
79 CTCGCAAGTCGGAGAGCGGCGAGGAACTGCTCTCTCTCAAGAGTGGAAAGCGCGTACCTG 138
62 CTCGCAAGTCGGAGAGCGGCGAGGAACTGCTCTCTCTCAAGAGTGGAAAGCGCGTACCTG 121
139 GAGCTGGAGAGTGGAGAGCGGCGAGGAACTGCTCTCTCTCTCAAGAGTGGAAAGCGCGTAC 198
122 GAGCTGGAGAGTGGAGAGCGGCGAGGAACTGCTCTCTCTCTCAAGAGTGGAAAGCGCGTAC 181
199 CCTCGATGTCAGAGTGGAAAGCGGCGAGGAACTGCTCTCTCTCTCAAGAGTGGAAAGCGCG 258
182 CCTCGATGTCAGAGTGGAAAGCGGCGAGGAACTGCTCTCTCTCTCAAGAGTGGAAAGCGCG 241
259 CCATTCCTTCCGCAAGTGGAGAGTGGAAAGCGGCGAGGAACTGCTCTCTCTCTCAAGAGT 318
242 CCATTCCTTCCGCAAGTGGAGAGTGGAAAGCGGCGAGGAACTGCTCTCTCTCTCAAGAGT 301
319 GTGCGAGAGAGTGGAGAGTGGAAAGCGGCGAGGAACTGCTCTCTCTCTCAAGAGTGGAAAG 378
302 GTGCGAGAGAGTGGAGAGTGGAAAGCGGCGAGGAACTGCTCTCTCTCTCAAGAGTGGAAAG 361
379 GGT 438
362 GGT 421
439 AGAGAGT 498
422 AGAGAGT 481
499 GTTTGGGAAATTTCTCAATTAAGATTAATTTGTAAATTTGCTTCTTCTTCTTCTTCTTCT 558
482 GTTTGGGAAATTTCTCAATTAAGATTAATTTGTAAATTTGCTTCTTCTTCTTCTTCTTCT 541
559 TGT 618
542 TGT 601

Qy 619 TCTGTACCCAGCACT---CGAACATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 675
Db 602 TGTGTACCCAGCACTGT 661
Qy 676 TGAAGGTTAAATAATAGATTAACGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 735
Db 662 TGAAGGTTAAATAATAGATTAACGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 721

RESULT 20

BY707632

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

MUS MUSCULUS

721 bp mRNA linear EST 16-DEC-2002
musculus cDNA clone 1810023A07 5', mRNA sequence.
BY707632.1 GI:27118809
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 721)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A.,
Quackenbush, J., Schirini, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbett,
L. E., Cousins, S., Daila, E., Dragan, T. A., Fletcher, C. E., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lennard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Mikl,
H., Nagashima, T., Numata, K., Okido, T., Pavau, W. J., Petrea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Poutis, J. U., Qi, D., Ring,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sanderlin, A., Schneider, C., Semple, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yangisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED
22354683
12466851
TITLE
CONTACT: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,
I., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,
Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono,
H., Koyama, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,
Numazaki, R., Ohno, K., Shibata, K., Shiraki, T., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,
Waki, K., Watanabe, A., Watanabe, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source location/Qualifiers
 1..721
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="1810023407"
 /sex="male"
 /tissue_type="pancreas"
 /dev_stage="10 day old"
 /clone_lib="RIKEN full-length enriched, 10 day old male pancreas"

BASE COUNT

190 a 147 c 200 g 184 t

ORIGIN

Query Match 60.3%; Score 687.2; DB 14; Length 721;
 Best Local Similarity 99.0%; Pred. No. 1e-136;
 Matches 713; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

19 GGGCCGACGTGGAGAGCGGAGGAGACCTGCGCTCTTCTTCCGACATCCGGAGCGCAG 78
 2 GGCCTGCGTGGAGAGCGGAGGAGACCTGCGCTCTTCTTCCGACATCCGGAGCGCAG 61
 79 CTCGAAGTCGGAGAGCGGAGGAGACCTGCGCTCTTCTTCCGAGAGGAGCGCATG 138
 62 CTCGAAGTCGGAGAGCGGAGGAGACCTGCGCTCTTCTTCCGAGAGGAGCGCATG 120
 139 GAGCTGGAGCGTTGAGTGCATACCTGTCACCTGTCAGAGGTCACAGTATGATGCTCG 198
 121 GAGCTGGAGCGTTGAGTGCATACCTGTCACCTGTCAGAGGTCACAGTATGATGCTCG 180
 199 CCTTCATGTCAAGCTGAAAAACAAGCAAGCAAGTGTGTTGTTGCTGGGGAGAGTTAA 258
 181 CCTTCATGTCAAGCTGAAAAACAAGCAAGCAAGTGTGTTGTTGCTGGGGAGAGTTAA 240
 259 CCATTCTTCCCAACTGCTGATGTCCTGCTGGGGAGAGCAAGCAATGCGTGCCTCT 318
 241 CCATTCTTCCCAACTGCTGATGTCCTGCTGGGGAGAGCAAGCAATGCGTGCCTCT 300
 319 GTGCCAGCAGACTGGTGTAGTCCAAAGATCGCAATGAGAGTGGCCAGCGCTCCT 378
 301 GTGCCAGCAGACTGGTGTAGTCCAAAGATCGCAATGAGAGTGGCCAGCGCTCCT 360
 379 GGTGTGTTGCTGACCTCTGCAACAAGACTPAACACTGCAAGGGATTCATCTTTAGAGAG 438
 361 GGTGTGTTGCTGACCTCTGCAACAAGACTPAACACTGCAAGGGATTCATCTTTAGAGAG 420
 439 AGAGGATGCTGTGCGCTTGTGAGACTCACCAAGGCTTCTTATTAATTTGCTGTTTA 498
 421 AGAGGATGCTGTGCGCTTGTGAGACTCACCAAGGCTTCTTATTAATTTGCTGTTTA 480
 499 GTTTGGGAAATTTCTTACAAATTAATTTGTTAAAAATGCGCTTCTTACCTCTG 558
 481 GTTTGGGAAATTTCTTACAAATTAATTTGTTAAAAATGCGCTTCTTACCTCTG 540
 559 TGTGTGTGTGTATAGCAATGATAGAGAGCGAAGACCAAGAAATATATCTTTGTTTA 618
 541 TGTGTGTGTGTATAGCAATGATAGAGAGCGAAGACCAAGAAATATATCTTTGTTTA 600

QY 619 TCTGACCCAGCACT---GGAACATTGTTTACAGAGAACATGTTGTTATGCT 675
 DB 601 TCTGACCCAGCACTGTGGAACTTGTGTTTACAGAGAACATGTTGTTATGCT 660
 QY 676 TGAGGTTTAAAAATAGATTAACGAATGTTCACGTACCAATTAATGATTTGAAAGCC 735
 DB 661 TGAGGTTTAAAAATAGATTAACGAATGTTCACGTACCAATTAATGATTTGAAAGCC 720

RESULT 21

LOCUS B1647378 721 bp mRNA linear EST 12-SEP-2001
 DEFINITION 603279634F1_NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5319872 5', mRNA sequence.

ACCESSION B1647378
 VERSION B1647378.1 GI:15561614

KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 731)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:

<http://image.llnl.gov>
 Plate: LLM11811 row: C column: 09

High quality sequence stop: 731.

FEATURES

source

1..731

location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/strain="129, C57BL/6J, FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:5319872"

/tissue_type="tumor, gross tissue"

/dev_stage="10 months"

/lab_host="DH10B"

/clone_lib="NCI CGAP Mam3"

/note="Organ: mammary; Vector: PCMV-SPOrt6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT

192 a 157 c 200 g 182 t

ORIGIN

Query Match 60.2%; Score 686.6; DB 12; Length 731;
 Best Local Similarity 98.8%; Pred. No. 1.3e-136;
 Matches 724; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 21 CCGAGCTGAGAGAGCGGAGGAAACCTGCGTCTTCTTCCGACCTCGGAGCGGAGGT 80
 DB 1 CCGAGCTGAGAGAGCGGAGGAAACCTGCGTCTTCTTCCGACCTCGGAGCGGAGGT 60
 QY 81 CCAAGTCGGAGAGCGCAAGATGTTCTCTCAAGATGGAACCGGTAGCCATGTGA 140
 DB 61 CCAAGTCGGAGAGCGCAAGATGTTCTCTCAAGATGGAACCGGTAGCCATGTGA 120
 QY 141 GCTGGAGCTTGAATGCGATACCTGTGCAATCTGAGAGGTCCAGGTATGATGCTGCC 200
 DB 121 GCTGGAGCTTGAATGCGATACCTGTGCAATCTGAGAGGTCCAGGTATGATGCTGCC 180

OY	201	TTGCATGTCGAAGCTGAAAACAAGCAAGAGAGACTGGTGTGTGGTCTGGGGAGAGGTAAAC	260
Db	181	TTGCATGTCAAGCTGAAAACAAGCAAGAGAGACTGGTGTGTGGTCTGGGGAGAGGTAAAC	240
OY	261	ATTCCTTCCACAACTGCTGCAATGTCCTGTGGGTGAAAACAGAACAAATGCCTCTGT	320
Db	241	ATTCCTTCCACAACTGCTGCAATGTCCTGTGGGTGAAAACAGAACAAATGCCTCTGT	300
OY	321	GCCAGCAGAGACTGGGTAGTCCAAAGATTCGGCAATGAGAGGTGGCCAGGCCCTCTGG	380
Db	301	GCCAGCAGAGACTGGGTAGTCCAAAGATTCGGCAATGAGAGGTGGCCAGGCCCTCTGG	360
OY	381	TGTGGTGTCTGACCCCTGGACAAAGACTAAACACTGCAGGGATTCATCTTGAAGAGAG	440
Db	361	TGTGGTGTCTGACCCCTGGACAAAGACTAAACACTGCAGGGATTCATCTTGAAGAGAG	420
OY	441	AGGATGCTGTGCGCCCTTGAAGCTCACCAAAAGCTTGTATTAAATTTGTCTGTTAGT	500
Db	421	AGGATGCTGTGCGCCCTTGAAGCTCACCAAAAGCTTGTATTAAATTTGTCTGTTAGT	480
OY	501	TTTGGGAAATTTCTTCACAATTAGTAATTTGTTTAAAAATGGCCCTTCTTACCTCTGGTG	560
Db	481	TTTGGGAAATTTCTTCACAATTAGTAATTTGTTTAAAAATGGCCCTTCTTACCTCTGGTG	540
OY	561	TGTGTGTGTGAATGCAATGCATAGAAAGCCGGAACACACAGAAATAGATCTTGTATTAC	620
Db	541	TGTGTGTGTGAATGCAATGCATAGAAAGCCGGAACACACAGAAATAGATCTTGTATTAC	600
OY	621	TGTACCCACGACT--GGAACAATTGTGTTCACAGAAAGACATTTGTGTATTAGCTTG	677
Db	601	TGTACCCACGACTGTGGGMACTGTGTGTTCACAGAAAGCA-TGTACAGTATATAGCTTG	659
OY	678	AGGTTTAAAAATAGATAACGAATGTTTACAGTAACTAAATAAATAGCATTTGAAAAGCCGA	737
Db	660	AGGTTTAAAAATAGATAACGAATGTTTACAGTAAAC-AATAAAATGCAITGAAAAGCCGA	718
OY	738	CTCCTCTTAATCC	750
Db	719	CTCCTCTTAATCC	731

```

RESULT 22
LOCUS      CB196142
DEFINITION CB196142      889 bp      mRNA      linear      EST 05-FEB-2003
            AGENCOURT_11258914 NIH-MGC_135 Mus musculus cDNA clone
IMAGE:0137902 5', mRNA sequence.
ACCESSION  CB196142
VERSION     CB196142
KEYWORDS   CB196142.1 GI:28223462
SOURCE     EST.
           Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 889)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgaabs@mail.nih.gov
            Tissue Procurement: Dr. David Rowe
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM0039 row: n column: 23
            High quality sequence scop: 718.
            Location/Qualifiers
                1..889
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
FEATURES
    source

```

```

/clone="IMAGE:30137902"
/lab host="DH10B (phage-resistant)"
/clone_lib="NH MGC 135"
/note="Vector: PCMVSPORT6.1, Site 1: EcoRV; Site 2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic 11mdb, maxilla and mandible, day 12.5, 13.5, 14.5,
and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dT. cDNA
Enrichment: 51k bp, Average insert size 1.6k bp.
Normalization (Cot value): 7.5 kb. Priming sequence:
5'-GACTAGTCTTAGATCGGAGCGGCGGCC(T)3'. Tissue contributed by
'David Rowe. Library constructed by Resesen, Invitrogen
Corp."

```

BASE COUNT	246 a	153 c	195 g	270 t	25 others
ORIGIN					

	Query Match	Similarity	59.7%;	Score	682.9;	DB	14;	Length	889;	
	Best Local	Similarity	98.7%;	Pred.	No. 8.8e-16;					
	Matches	752;	Conservative	0;	Mismatches	2;	Indels	8;	Gaps	6;
OY	357	GAGAGTGCGCCGACGGCCTC-	CTGGTGTGTTGTCGACCCTTGAGCAAAAGACTMAA	CAC	TG	415				
Db	1	GAGAGTGGCCGACGCCTCTC	CGGTGTGGTGTGACCCCTGGACAAMAGACTMAAAC	CTG		60				
OY	416	CAGGGATTCANCCCTGAGAGAGAGA	ATGCTGAGGCCCTTGAGACTCACCAAAGCT			475				
Db	61	CAGGGATTCAACCTTGAGAGAGAGAG	ATGTGTGCGCTTTGAGACTCACCAAAGCT			120				
OY	476	TGCTTTATAATTGTCTGTGTAGTTT	TGGAAATTCCTCAAAATTAAGATTAATTGTTA			535				
Db	121	TGCTTTATAATTGTCTGTGTAGTTT	TGGGAATTCCTCAATTAAGATTAATTGTTA			180				
OY	536	AAATGCGCTTTCCTCACTCTGTGTGT	GTGTGTGTGTATCGAATGATAGGAAGGAGAA			595				
Db	181	AAAATGCGCTTTCCTCACTCTGTGTGT	GTGTGTGTATCGAATGATAGGAAGGAGAA			240				
OY	596	CACGAGAATAATGATTTGTTATCTGA	CCACGACACT---GGAATGTGTCTACAGA			652				
Db	241	CACGAGAATAATGATTTGTTATCTGA	CCACGACACTGTGGGAACATTTGTTCA	CGA		300				
OY	653	AGAACATTGTTGTGTTATGCTTGAGG	GGTTAAAAATATGATAACGAATGTTACATRA			712				
Db	301	AGAACATTGTTGTGTTATGCTTGAGG	GGTTAAAAATATGATAACGAATGTTACATRA			360				
OY	713	CAAAATTAATGATTAAGAAAGCCGAC	TCCTCCTATCCTTTTGTGTTGGAGAGAGCA			772				
Db	361	CAAAATTAATGATTAAGAAAGCCGAC	TCCTCCTATCCTTTTGTGTTGGAGAGAGCA			420				
OY	773	AGCGAGCGCACCCCTGCTCTTCA	TTTGCTGTGAATAGAGATTTTAACTCGACTCA	GTG		832				
Db	421	AGCGAGCGCACCCCTGCTCTTCA	TTTGCTGTGAATAGAGATTTTAACTCGACTCA	GTG		480				
OY	833	AAGAGCGGTAACTGTGCGGTAAAC	TGTATATATGCGGTAACTGTGCGGTAAAC	GGCTTTGT		892				
Db	481	AAGAGCGGTAACTGTGCGGTAAAC	TGTATATATGCGGTAACTGTGCGGTAAAC	GGCTTTGT		540				
OY	893	CTCCGACTTCCCANCTTGACTTGGC	CAGAGAGCCTGATTTGTCCAACCACTTAGTTC			952				
Db	541	CTCCGACTTCCCANCTTGACTTGGC	CAGAGAGCCTGATTTGTCCAACCACTTAGTTC			600				
OY	953	TAAAGACTGTTTCTGTGTTTTGCC	GGAAGTGTATATGTATGTTTATGTC	CAAAATATTA		101				
Db	601	TAAAGACTGTTTCTGTGTTTTG	-CGAAGTGTATATGTATGTTTATGTC	-AAAAATATA		658				
OY	1013	GTAGAGAAATGCGCTTATCTATTA	CACTGAAGTTCAATATGTCATATGTTTAATAATA			107				
Db	659	GTAGAGAAAT-CTTACTATGATTA	CACTGAAGTTCAATATGTCATATGTTTAATAATA			717				
OY	1073	TTGTGC-TTTGAGTATTAAGTTGA	CTATATCTTTAAAA	1113						
Db	718	TTGTGCTTTGAGTTTAAAGTTGA	TATATATCTTTAAAA	759						

LOCUS	CA464071	867 bp	mRNA	linear	EST 12-NOV-2002
DEFINITION	AGENCOURT_10722790 NIH_MGC_169 Mus musculus cDNA clone				
IMAGE	6771682 5', mRNA sequence.				
ACCESSION	CA464071				
VERSION	CA464071.1	GI:24920423			
KEYWORDS	EST.				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	Mus musculus				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (bases 1 to 867)				
JOURNAL	NIH-MGC http://mgc.ncbi.nlm.nih.gov/ .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgabbs-r@mail.nih.gov				
	Tissue Procurement: Dr. Jonathan Kuo, NIMH				
	cDNA Library Preparation: Michael Brownstein Laboratory				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov				
	Plate: L16CM3091 row: k column: 09				
	High quality sequence stop: 548.				
FEATURES	Location/Qualifiers				
source	1..867				
	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/db_xref="taxon:10090"				
	/clone="IMAGE:6771682"				
	/lab_host="DH10B (T1-phage-resistant)"				
	/clone_lib="NIH_MGC_169"				
	/note="Organ: Testicles; Vector: pDNR-LIB; Site_1: Sfil (ggccatcagcc); Site_2: Sfil (ggccgcctggcc); cDNA made by oligo-dt priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCATGTGATCAACGACGAGATGGCATTCGCGCGG-3' and 5'-ATTCTAGAGCGGACGCGGACGATG-3' (30)NN-3'. Full-length enriched library was constructed using the Clontech created SMART kit and size-selected to contain the 0.5 kb size fraction. Library created in the laboratory of M. Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."				
BASE COUNT	229 a	184 c	261 g	187 t	6 others
ORIGIN					
Query Match	59.7%	Score 680.2	DB 14	Length 867	
Best Local Similarity	93.8%	Pred. No. 3.2e-135			
Matches 738	Conservative 0	Mismatches 44	Indels 5	Gaps 3	
1	GTTCGCGCGCGCCGATGCGCGAGCTGGAGACGCGGAGAACCCGTCCTTCTTC	60			
23	GTTCGCGCGCGCCGATGCGCGAGCTGGAGACGCGGAGAACCCGTCCTTCTTC	82			
61	GCACTCCGCGGAGCGCAGGCTCCAAGTCGGAGGCGACAAGATGTCTCTCAAGAAGTG	120			
83	GCACTCCGCGGAGCGCAGGCTCCAAGTCGGAGGCGACAAGATGTCTCTCTCAAGAAGTG	142			
121	GAAAGCGGTAGCCATGTGGAGCTGGAGCGTTGAGTGCATCTGTGCGCATCTGAGAGGT	180			
143	GAAAGCGGTAGCCATGTGGAGCTGGAGCGTTGAGTGCATCTGTGCGCATCTGAGAGGT	202			
181	CCAGGTGATGATGCTGCGCTTTCATGTCAAGCTGAAAAACAAGCAAGAGACTGTGTGT	240			
203	CCAGGTGATGATGCTGCGCTTTCATGTCAAGCTGAAAAACAAGCAAGAGACTGTGTGT	262			
241	GGTCTGGGAGAGTGTATCAATCTCTTCAACAACCTGCGCATGTCTCTGGAGTGAACA	300			
263	GGTCTGGGAGAGTGTATCAATCTCTTCAACAACCTGCGCATGTCTCTGGAGTGAACA	322			
301	GAAACATCTGCGCTCTGTGCGACGACGAGCTGGTGTGTCACCAAGATGCGCAATGAGA	360			

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT																						
Db	323	GAACAAATGCGCCCTCTGTGACAGAGAGCTGGGTAATGCCAAATGGAATGGA	361	GGTGGCCAGGCGCTCCCTGGTGTGTGCTGACACCTTGAGAAAGACTGACGAGG	383	GGTGGCCAGGCGCTCCCTGGTGTGTGCTGACACCTTGAGAAAGACTGACGAGG	421	GATTCATCTTGAGAGAGAGATGCTGTGCGCTTTGAGACTCACCAAGGCTTGCT	443	GATTCATCTTGAGAGAGAGATGCTGTGCGCTTTGAGACTCACCAAGGCTTGCT	481	TATTAATTTGCTGTAGTTTGGGAAATTCCTCAATTAAGTAAATTTGTTAAAT	503	TATTAATTTGCTGTAGTTTGGGAAATTCCTCAATTAAGTAAATTTGTTAAAT	541	GGCCTTTCCTACCTCT-GGTGTGTGTGTGATACGAATCATGAGAGACGAGAACCC	563	GGCCTTTCCTACCTCTGGGTGTGTGTGTGTGATACGAATCATGAGAGACGAGAACCC	600	AGAAATGATCTTTGTTATCTGTACCCACGACT--GGAACATTTGTTCACAGAAAG	623	AGAAATGATCTTTGTTATCTGTACCCACGACTGGGGGACATTTGTTCACAGAAAG	657	CATT-GTTTGTTTATGCTTGAGAGGTTTAAATTAATTAAGCATGTTACGTAACA	683	CATTGTTTGTTTATGCTTGAGAGGTTTAAATTAATTAAGCATGTTACGTTACCA	716	ATAAATGCAATGAAAGCGACTCCTCTCATCTTTTGTGTGTGGAGAGAGAGCAAGC	743	ATAAATGCAATGAAAGCGACTCCTCTCATCTTTTGTGTGTGGAGAGAGAGCAAGC	776	GAGGCCA 782	803	GAGGNA 809
Db	803	GAGGNA 809	776	GAGGCCA 782	803	GAGGNA 809																												
RESULT 24	BI156184	602903365F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5032984 5', mRNA sequence.	BI156184	BI156184.1	GI:14616185	EST.	Mus musculus (house mouse)	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus. (bases 1 to 978)	NIH-MGC http://mgs.nci.nih.gov/.	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished	Contact: Robert Strausberg, Ph.D. Email: egadps-remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLAM1092 row: a column: 17 High quality sequence stop: 855.																					
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers																						
FEATURES	SOURCE	1..978	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	Location/Qual																											

```
/clone.lib="NCI CGAP Mem3"  
/note=Organ: mammary; Vector: pCMV-Sport6; site 1: SalI;  
site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lotzner Hemmighausen/Chu-Xia Deng, NIH  
Reference for transgenic model: Xu et al., Nature Genetics  
22, 37-43 (1999)."
```

BASE COUNT	250 a	199 c	260 g	26
ORIGIN				

Query Match	59.1%	Score	673.4	DB	12	Length	978
Best Local Similarity	90.5%	Pred	No. 9.1e-134				
Matches	837	Conservative	0	Mismatches	61	Indels	27
						Gaps	10

OY	75	CAGCGTCCAAAGTCGGGAGGCGCA	CAAGAATGTTCTCTCAAGAAAGTGGAACGGCGATGACA	134
Db	1	CAGGCTCCAAAGTCGGGAGGCGCA	AGAATGTTCTCTCTCAAGAAAGTGGAACGGCGATGACA	60
OY	135	TGTGAGAGCTGGAGACGTTGAGTGC	ATACCTGTGCCATCTGCAGGGTCCAGTGATGATG	194
Db	61	TGTGAGAGCTGGAGACGTTGAGTGC	ATACCTGTGCCATCTGCAGGGTCCAGTGATGATG	120
OY	195	CCTGCCCTTCAGATGTCAAGCTGAAA	ACAAGCAGAAGACCTGTGTGTGTGTGGGAGAGT	254
Db	121	CCTGCCCTTCAGATGTCAAGCTGAAA	ACAAGCAGAAGACCTGTGTGTGTGTGGGAGAGT	180
OY	255	GTAACCATTCCTTCCACAACACTCTGC	ATGTCCTGCTGGTGGTGAACAGAACATTCGTCGCC	314
Db	181	GTAACCATTCCTTCCACAACACTCTGC	ATGTCCTGCTGGTGGTGAACAGAACATTCGTCGCC	240
OY	315	CTCTGTGCAGACAGACCTGGGTATG	TCCAAAGATCGGCAATAGAGAGTGGCCCAAGGCC	374
Db	241	CTCTGTGCAGACAGACCTGGGTATG	TCCAAAGATCGGCAATAGAGAGTGGCCCAAGGCC	300
OY	375	TCTGTGTGTGTGTGTGTGTGTGTGT	GTGACCTGTGACCAAGACCTTAACCTGTGAG	434
Db	301	TCTGTGTGTGTGTGTGTGTGTGTGT	GTGACCTGTGACCAAGACCTTAACCTGTGAG	360
OY	435	AGAGAGAGAGATCTGTGGCCCTTTG	AGACTCACCAAAAGCCTTATTATTTGTCTG	494
Db	361	AGAGAGAGAGATCTGTGGCCCTTTG	AGACTCACCAAAAGCCTTATTATTTGTCTG	420
OY	495	TTTAAGTTTGGGAAATTCCTCAACA	ATAATGATTAATTTGTTAAAAATGGCCCTTCCACTC	554
Db	421	TTTAAGTTTGGGAAATTCCTCAACA	ATAATGATTAATTTGTTAAAAATGGCCCTTCCACTC	480
OY	555	CTGTGTGTGTGTGTGTATAGAA	TGATAGAAAGAGGAGAACACCAAGAAATGATCTTTG	614
Db	481	CTGTGTGTGTGTGTATAGAA	TGATAGAAAGAGGAGAACACCAAGAAATGATCTTTG	540
OY	615	TTTATCTGTACCCACGACT---	GGACATTGTGTTTCAAGAAACAATGTTGTGTTTA	671
Db	541	TTTATCTGTACCCACGACTGTGGAA	CATTGTGTTTCAAGAAACAATGTTGTGTTTA	600
OY	672	TGCTTGAAGGTTAAAAAATAGATAA	ACGATG-TTACAGTAAACAAATAAATCATGAA	730
Db	601	TGCTTGAAGGTTAAAAAATAGATAA	ACGATG-TTACAGTAAACAAATAAATCATGAA	660
OY	731	AAGCCGACTCTCTCTTAATCTTTT	TGTGTGGAGAGGCAAGCCAG----GCCACC	785
Db	661	AAGCCGACTCTCTCTTAATCTTTT	TGTGTGGAGAGGCAAGGGAAGGCCACCCCT	720
OY	786	TGCTGTCTTCAATTTGCT----	GTGAATGAGATTTTAAACCTGCACCTAGTGAAG-AGGC	839
Db	721	GTCTGTCTTCAATTTGCTTTGCTTT	GTGAATGAGAGGCAATTTTAAACCTGCACCTAGTGAAGGCC	780
OY	840	GTAACTGTGCGG--TTAACTGTAA	TATATGCGCTAA----CTGTGGGTTAAACGGCTTTGTC	893
Db	781	GTAACTGTGCGGGAATCTGTAA	TATATGCGCTAAACCTGTCCGGGTTAAACGGCTTTGTC	840
OY	894	T--CCTGACTTCTCACTCTT-----	GACTTGGCCAGGAAGGC-TGGAATGTTCAACACTT	947
Db	841	TCCTGACTTCTCACTCTTGTGA	ACTTTTGAGCCAGGAAGCCTTGGGAATGTCAACACTT	900

```

Qy      948 AGTCTAAGAAGACTGTTTCTGTTT 972
          |||||
Db      901 AGTCTAAGCAACCGGTTCTGTT 925

```

RESULT	25
CBS18729	
LOCUS	CBS18729 787 bp mRNA linear EST_04-MAR-2003
DEFINITION	AGNCOCURT_12403690 NIH_MGC_165 Mus musculus cDNA clone IMAGE:30285241 5' , mRNA sequence.
ACCESSION	CBS18729
VERSION	CBS18729.1 GI:28842964
KEYWORDS	EST,
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCE	1 (bases 1 to 787)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D. strausberg@nhi.nih.gov

FEATURES

Location/Qualifiers

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_image="30285241"
/tissue_type="Primary cultures of Sertoli cells"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH MGC 165"
/note="Organ: testis; Vector: pDNR-LIB; Site 1: SfiI
(gggcatatggcc); Site 2: SfiI (ggcgccctggcc); 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCCATTTATGGCC-3' and 3' adaptor sequence:
5'-ATTTCAGAGCGCGAGCGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4 kb
(range 0.6-3.5 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA).
Note: this is a NIH MGC library."

```

BASE COUNT	209 a	161 c	214 g	203 c
ORIGIN				

Query Match	Score	DB 14	Length
58.9%	671	787	

Best Local Similarity 95.8%; Pred. NO. 3e-153;
Matches 723; Conservative 0; Mismatches 25; Indels 7; Gaps 3

68 GGGAGCCGACGCTCCAAGTCGGAGCGACAAGATGTTCTCTCAAGAAGTGAACGG 127

Db 4 GGGAGCGCAGGCTCCAGTCCGGAGCGACAAGTGTCTCTCTCAAGAACTGGAACGG 63

BASE COUNT	209 a	161 c	214 g	203 t	
ORIGIN					
Query Match		58.9%;	Score 671;	DB 14;	Length 787;
Best Local Similarity		95.8%;	Pred. No. 36-133;		
Matches 723;	Conservative	0;	Mismatches 25;	Indels	7; Gaps
					3
QY	68	GGGAGCGCGAGGCTCCAA	GTCCGAGGCGCA	CAAGATGTTCTCTCAAGAGTGGAA	CGCG 127
DB	4	GGGAGCGCGAGGCTCCAA	GTCCGAGGCGCA	CAAGATGTTCTCTCAAGAGTGGAA	CGCG 63
QY	128	GTAGCCATGTGAGAGCTGTG	GACGTTGAGTCCGATCTGTC	CGCATCTGCAAGGCTCCAGG	GTG 187
DB	64	GTHGCCATGTGAGAGCTGTG	GACGTTGAGTCCGATCTGTC	CGCATCTGCAAGGCTCCAGG	GTG 123
QY	188	ATGAGATGCTGCGCTTCGAT	GTCAAGCTGAAACCAAGCA	AAGAAGACTGTGTTGTGGCTGG	247
DB	124	ATGAGATGCTGCGCTTCGAT	GTCAAGCTGAAACCAAGCA	AAGAAGACTGTGTGGCTGG	183
QY	248	GGAGAGTGTAAACATCTTC	CAACAACCTGTCATCTCC	TGTGGGTGAAACGAACAT	307


```

Db      184 GGAAGTGAACCAATCTTCCACAACTGCTGATCTCCCTGTGGTGAAGCAAGCAAT 243
QY      308 CGCTGCCCTCTGTGTCCAGCAGACTGGGTAGTCCAAAGATCGCAAAATGAGAGTGCC 367
Db      244 CGCTGCCCTCTGTGTCCAGCAGACTGGGTAGTCCAAAGATCGCAAAATGAGAGTGCC 303
QY      368 CAGGCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 427
Db      304 CAGGCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 363
QY      428 CTTTGAAGAGAGAGATGCTGTGCGCTTTGAGACTCAACAAAGCTTGTATTAT 487
Db      364 CTTTGAAGAGAGAGATGCTGTGCGCTTTGAGACTCAACAAAGCTTGTATTAT 423
QY      488 TTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 547
Db      424 TTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483
QY      548 CCTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 607
Db      484 CCTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 543
QY      608 ATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 664
Db      544 ATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 603
QY      665 GTTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 724
Db      604 GTTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 663
QY      725 ATTGAAGACCGACTCTCTCTTAATCCTTTTGTG--TTGGAGAGAGCGAAG-CGAGCG 780
Db      664 ATTGAAGACCGACTCTCTCTTAATCCTTTTGTG--TTGGAGAGAGCGAAG-CGAGCG 723
QY      781 CACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 815
Db      724 CACCCCGGCTGTCTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 758

```

RESULT 26

```

Bg917948      916 bp      mRNA      linear      EST 05-JUN-2001
LOCUS      602820877F1 NCI_CGAP_Mus musculus cDNA clone IMAGE:4949851 5'
DEFINITION      mRNA sequence.
ACCESSION      Bg917948
VERSION      Bg917948.1 GI:14298424
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 916)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10903 row: 1 column: 20
High quality sequence stop: 760.
Location/Qualifiers
1..916

```

```

FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4949851"

```

```

/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mus"
/organism="Mus musculus"
/site_2="NCI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      232 a      202 c      264 g      218 t
ORIGIN

```

```

Query Match      58.4%; Score 665.4; DB 12; Length 916;
Best Local Similarity 93.8%; Pred. No. 4.7e-132;
Matches 761; Conservative 0; Mismatches 36; Indels 14; Gaps 6;

```

```

QY      1 GTTTCGCCGCCGCCGCATGCGAGCTGGAGAGCGCGAGAACCTGCGTCTTCTTC 60
Db      40 GTTTCGCCGCCGCCGCATGCGAGCTGGAGAGCGCGAGAACCTGCGTCTTCTTC 99
QY      61 GCACCTCCGGAGAGCGAGCTCCAAAGTCGGAGAGCGACAAGATGTTCTCTCAAGAGT 120
Db      100 GCACCTCCGGAGAGCGAGCTCCAAAGTCGGAGAGCGACAAGATGTTCTCTCAAGAGT 159
QY      121 GAACCGGTAGCCATGTGAGCTGGAGCTTGTAGTCCATCTGTGCATCTGAGAGT 180
Db      160 GAACCGGTAGCCATGTGAGCTGGAGCTTGTAGTCCATCTGTGCATCTGAGAGT 219
QY      181 CAGGTGATGATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db      220 CAGGTGATGATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 279
QY      241 GGTCTGGGGAGAGGTAAACCATCTCTCCACAACTGTGCAATGCTGCTGGGTGAAAGA 300
Db      280 GGTCTGGGGAGAGGTAAACCATCTCTCCACAACTGTGCAATGCTGCTGGGTGAAAGA 339
QY      301 GAACATGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db      340 GAACATGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 399
QY      361 GGTGCCCAAGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Db      400 GGTGCCCAAGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 459
QY      421 GATTATCTTTAAGAGAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db      460 GATTATCTTTAAGAGAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519
QY      481 TATTATTTGTCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db      520 TATTATTTGTCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579
QY      541 GGCCTTTCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db      580 GGCCTTTCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 638
QY      601 GAAATGATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 657
Db      639 GAAATGATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 698
QY      658 ATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 716
Db      699 CATG--TGGGTTATGCTTGAAGGTTAAATAATGATTAAGATGTTCAGTAAACA 756
QY      717 TAAAA--TGCATGAAAAGCGACTCTCTCAATCTTTTGTGT--TGGAGAGAG 769
Db      757 ATTTAAATGAGCGTGAAGAAGCGAATCTCTCAATCGTTGTGTGTGTGTGTGTGTGTGT 816
QY      770 GGAAGCGAGCGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 800
Db      817 GGAAGCGAGCGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 847

```

```

RESULT 27
CA977674          944 bp  mRNA  linear  EST 06-JAN-2003
LOCUS             CA977674
DEFINITION       IMAGE:30146934 5', mRNA sequence.
ACCESSION        CA977674
VERSION          CA977674.1 GI:27510328
KEYWORDS         EST.
SOURCE           Mus musculus (house mouse)
ORGANISM         Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE        NIH-MGC http://mgi.nci.nih.gov/
AUTHORS          NIH-MGC Consortium (LNL)
TITLE            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL          Unpublished
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-remail.nih.gov
                  Tissue Procurement: Dr. David Rowe and Dr. Mina
                  CDNA Library Preparation: Invitrogen Corp
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
                  DNA Sequencing by: Agencourt Bioscience Corporation
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LNL at:
                  http://image.llnl.gov
                  Plate: NDAM0063 row: 9 column: 07
                  High quality sequence stop: 673.

FEATURES
    source
        1..944
            /organism="Mus musculus"
            /mol_type="mRNA"
            /db_xref="taxon:10090"
            /clone="IMAGE:30146934"
            /lab_host="DH10B (phage-resistant)"
            /note="Vector: pCMV-Sport6.1, ccdB; Site 1: EcoRV; Site 2:
            NOTI; Non-normalized full-length enriched library from
            pooled mouse embryonic limb, maxilla and mandible, day
            10.5 and 11.5 (size selected for the 0.5-1 kb fragments)
            Cloned directionally, priming method: Oligo-dT. CDNA
            enrichment: >1k bp. Average insert size 1.8k bp. Priming
            sequence: 5'-GACTAGTCTGATCGGAGCGGCCCT(T) 3'. Tissue
            contributed by, David Rowe. Library constructed by Resgen,
            Invitrogen Corp."
BASE COUNT       216 a 162 c 242 g 208 t 116 others
ORIGIN
Query Match      58.3%; Score 664.6; DB 14; Length 944;
Best Local Similarity 86.8%; Pred. No. 7e-132;
Matches 683; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

QY 21 CCGACGTGAGAGACGGCGAGGAACCTGCGCTCTTCTTGACATCCGGAGCGCAGGCT 80
DB 1 CCGACGTGAGAGACGGCGAGGAACCTGCGCTCTTCTTGACATCCGGAGCGCAGGCT 60

QY 81 CCAAGTGGGAGGCGCAAGATGTTCTCTCAAGAGTGAACGGGTGACATGTGA 140
DB 61 CCAAGTGGGAGGCGCAAGATGTTCTCTCAAGAGTGAACGGGTGACATGTGA 120

QY 141 GCTGGGAGCTTGAATGGGATCTGCTGTCATCTGCGAGGTCCAGGTGATGCTGCC 200
DB 121 GCTGGGAGCTTGAATGGGATCTGCTGTCATCTGCGAGGTCCAGGTGATGCTGCC 180

QY 201 TTGCATGTCAAGCTGAAAAACAAGACAGAGACTGTGTGTGTGTGGGAGAGTGAAC 260
DB 181 TTGCATGTCAAGCTGAAAAACAAGACAGAGACTGTGTGTGTGTGGGAGAGTGAAC 240

QY 261 ATTCTCTTCCAACTGCTGCAATGTCCTGTGGGAGGAACAAGCAATCGCTGCTGT 320
DB 241 ATTCTCTTCCAACTGCTGCAATGTCCTGTGGGAGGAACAAGCAATCGCTGCTGT 300

QY 321 GCCAGCAGGACTGGGTAGTCCAAAGAAATCGGCAATGAGAGTGGCCAGGCGCTCTGG 380
DB 301 GCCAGCAGGACTGGGTAGTCCAAAGAAATCGGCAATGAGAGTGGCCAGGCGCTCTGG 360

```

```

DB 301 GCCAGCAGGACTGGGTAGTCCAAAGAAATCGGCAATGAGAGTGGCCAGGCGCTCTGG 360
QY 381 TGTGTTGCTGACCTCTGACCAAAAGACTAAACCTGCAAGGGATTCATCTTGAGAGAG 440
DB 361 TGTGTTGCTGACCTCTGACCAAAAGACTAAACCTGCAAGGGATTCATCTTGAGAGAG 420
QY 441 AGATGCTGTGGCCCTTTGAGACTCAACAAAGCTGCTTATTAATTTGTCCTTACG 500
DB 421 AGATGCTGTGGCCCTTTGAGACTCAACAAAGCTGCTTATTAATTTGTCCTTACG 480
QY 501 TTGGGAAATTTCTCAATTAAGATAATTTGTTAAATAGGCTTCTTACCTGCTG 560
DB 481 TTGGGAAATTTCTCAATTAAGATAATTTGTTAAATAGGCTTCTTACCTGCTG 540
QY 561 TGTGTGTGATACGAATGATGAAGAGGAGAAACACCGAATAAATCTTGTATC 620
DB 541 TGTGTGTGATACGAATGATGAAGAGGAGAAACACCGAATAAATCTTGTATC 600
QY 621 TGTACCCAGACT---GGAACATTTGTTCACAGAGACATTTGTTATGCTTG 677
DB 601 TGTACCCAGACTGTGGGAAATTTGTTTCAAGAGAACATTTGTTATGCTTG 660
QY 678 AGGTTAAAAATAGATTAACGAATGTTACAGTAACAAATAAATGCAATTAAGCCGA 737
DB 661 AGGTTAAAAATAGATTAACGAATGTTACAGTAACAAATAAATGCAATTAAGCCGA 720
QY 738 CTCTCTCTAATCTTTTGTGTTGGGAGAGAGCAAGCGACCCCTGCTGCTCAT 797
DB 721 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780
QY 798 TTGCTGT 804
DB 781 TTTTAT 787

```

```

RESULT 28
BO177475          676 bp  mRNA  linear  EST 30-APR-2002
LOCUS             BO177475
DEFINITION       UT-M-ER0-bwo-n-12-0-UT-rl NIH BMP2_ER0 Mus musculus cDNA clone
ACCESSION        BO177475
VERSION          BO177475.1 GI:20352967
KEYWORDS         EST.
SOURCE           Mus musculus (house mouse)
ORGANISM         Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE        NIH-MGC http://mgi.nci.nih.gov/
AUTHORS          NIH-MGC Consortium (LNL)
TITLE            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL          Unpublished
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-remail.nih.gov
                  Tissue Procurement: Dr. James Lin, University of Iowa
                  CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
                  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LNL at:
                  http://image.llnl.gov
                  This clone was contributed by the Brain Molecular Anatomy Project
                  (BMAP)
                  Seq primer: pYX-5.

FEATURES
    source
        1..676
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="CS7BL/6"
            /db_xref="taxon:10090"
            /clone="IMAGE: 5699915"
            /tissue_type="whole brain"
            /dev_stage="embryo 15.5 dpc"
            /lab_host="DH10B (T1 phage resistant)"

```

/lone lib="NIH_BMAP_ER0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GTGGCTGGA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 167 a 143 c 192 g 174 t
 ORIGIN
 Query Match 57.3%; Score 653.4; DB 13; Length 676;
 Best Local Similarity 99.4%; Pred. No. 1.7e-129;
 Matches 667; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 QY 21 CCGACGTGAGGAGCGCGAGGAACTCTGCTCTTCTTCCGACCTCCGGAGCGAGGCT 80
 DB 1 CCGACGTGAGGAGCGCGAGGAACTCTGCTCTTCTTCCGACCTCCGGAGCGAGGCT 60
 QY 81 CCAAGTCGGAGGCGCAAGAGTGTCTCTCTCAAGAGTGAACCGCGTACCATGTGA 140
 DB 61 CCAAGTCGGAGGCGCAAGAGTGTCTCTCTCAAGAGTGAACCGCGTACCATGTGA 120
 QY 141 GCTGGAGCGTGTAGTSCGATACCTGTCATCTGCGAGGCGCAGGTGATGATGCTGCG 200
 DB 121 GCTGGAGCGTGTAGTSCGATACCTGTCATCTGCGAGGCGCAGGTGATGATGCTGCG 180
 QY 201 TTTCGATGTCAAGCTGAAAAACAAGCAAGAGAGCTGTGTGTCTGTGGAGAGTGTAC 260
 DB 181 TTTCGATGTCAAGCTGAAAAACAAGCAAGAGAGCTGTGTGTGTGTGGAGAGTGTAC 240
 QY 261 ATTCTCTCCACAACCTGCTGATGTCCCTGTGGTGAACAAGCAATCCGCTCTGT 320
 DB 241 ATTCTCTCCACAACCTGCTGATGTCCCTGTGGTGAACAAGCAATCCGCTCTGT 300
 QY 321 GCCAGAGAGCTGGGTGATCCAAAGAAATGGCAATGAGAGGCGCCAGGCGCTCTGG 380
 DB 301 GCCAGAGAGCTGGGTGATCCAAAGAAATGAGAGGCGCCAGGCGCTCTGG 360
 QY 381 TGTGTGTGCTGACCTCTGACCAAGACTAACTGCGAGGAGATTCTCTTGAAGAGAG 440
 DB 361 TGTGTGTGCTGACCTCTGACCAAGACTAACTGCGAGGAGATTCTCTTGAAGAGAG 420
 QY 441 AGGATGCTGTGGCGCTTTGAGACTACCAAGAGCTGTTTATTAATTTGTCTGTAGT 500
 DB 421 AGGATGCTGTGGCGCTTTGAGACTACCAAGAGCTGTTTATTAATTTGTCTGTAGT 480
 QY 501 TTGGGAAATTCCTCAATTAAGATAATTTGTTAAATTTGGCTTCTTACTCTGTGG 560
 DB 481 TTGGGAAATTCCTCAATTAAGATAATTTGTTAAATTTGGCTTCTTACTCTGTGG 540
 QY 561 TGTGTGTGATACGATGATGAGAGAGAGAGACCAAGAAATGATCTTTGTTATTC 620
 DB 541 TGTGTGTGATACGATGATGAGAGAGAGAGACCAAGAAATGATCTTTGTTATTC 600
 QY 621 TGTACCCAGCACT--GAAACATTTGTGTTCAGAAACATTTGTTGTTATGCTTG 677
 DB 601 TGTACCCAGCACTGTTGGAAACATTTGTTTTCAGAAACATTTGTTGTTATGCTTG 660
 QY 678 AGGGTTAAAA 688
 DB 661 ACGGTTAAAA 671

RESULT 29

BQ044727
 LOCUS BQ044727 686 bp mRNA linear EST 28-MAR-2002
 DEFINITION U1-M-EMO-bwm-j-14-0-UI.r1 NIH_BMAP_EMO Mus musculus cDNA clone
 IMAGE:5694061 5', mRNA sequence.
 ACCESSION BQ044727
 VERSION BQ044727.1 GI:19795492
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 686)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..686
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5694061"
 /tissue_type="whole brain"
 /dev_stage="embryo 18.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EMO"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GAGCAACGAC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institute of Mental Health
 (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 179 a 135 c 174 g 197 t
 ORIGIN
 Query Match 57.2%; Score 651.6; DB 12; Length 686;
 Best Local Similarity 98.7%; Pred. No. 4.2e-129;
 Matches 678; Conservative 0; Mismatches 5; Indels 4; Gaps 2;
 QY 231 ACTGTGTGTGCTGGGAGAGTGAACCATCTCTCCACAACGCTGATGCTCTGT 290
 DB 1 ACTGTGTGTGCTGGGAGAGTGAACCATCTCTCCACAACGCTGATGCTCTGT 60
 QY 291 GGGTAAACAGAAACATGCTGCTCTGTGCGACAGAGCTGGTAGTCCAAAGATG 350
 DB 61 GGGTAAACAGAAACATGCTGCTCTGTGCGACAGAGCTGGTAGTCCAAAGATG 120
 QY 351 GCAATGAGAGGTGGCGCCAGGCGCTCTGTGTGTGTGCTGACCTCTGCAAGACTAA 410
 DB 121 GCAATGAGAGGTGGCGCCAGGCGCTCTGTGTGTGTGCTGACCTCTGCAAGACTAA 180
 QY 411 CACTGAGGAGATCATCTTGAAGAGAGAGAGATGTGTGTGCTTGAAGACTCA 470


```

QY 452 CGCCTTGAGACTCACCAGAGCTGCTTATTAATTTGCTTTAGTTTGGAAAT 511
    |||||
Db 362 CGCCTTGAGACTCACCAGAGCTGCTTATTAATTTGCTTTAGTTTGGAAAT 421
QY 512 CTCCTCAATTAAGATTAATTTGTTAAAAATGCGCTTCTCACTGCTGTGTGTGA 571
    |||||
Db 422 CTCCTCAATTAAGATTAATTTGTTAAAAATGCGCTTCTCACTGCTGTGTGTGA 481
QY 572 TACGATGATAGAGAGGAGAACCCGAAATATCTTTGTTATCTGACCCAGA 631
    |||||
Db 482 TACGATGATAGAGAGGAGAACCCGAAATATCTTTGTTATCTGACCCAGA 541
QY 632 CT---GGAATTTGTTGACAGAGAAATGTTGTTTATGCTTGAAGGTTAAAA 688
    |||||
Db 542 CTGTGGAAACATTTGTTTTCACAGAGAACATTTGTTTATGCTTGAAGGTTAAAA 601
QY 689 ATAGATTAACGATGTTTACAGTAAACAATTAATGATTAAGGCGACTCTCTTAAT 748
    |||||
Db 602 ATAGATTAACGATGTTTACAGTAAACAATTAATGATTAAGGCGACTCTCTTAAT 661
QY 749 CCTTTT 755
    |||||
Db 662 CCTTTT 668

```

```

RESULT 34
BI556918 737 bp mRNA linear EST 05-SEP-2001
LOCUS 60339970F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5292832 5'
DEFINITION
B1556918
ACCESSION
B1556918
VERSION
B1556918.1 GI:15444232
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 737)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM11740 row: 1 column: 17
High quality sequence scop: 685.
Location/Qualifiers
1. 737
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5292832"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_1lb="NCI_CGAP_Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT,
library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
```

BASE COUNT

200 a 152 c 212 g 173 t

```

Query Match 56.7%; Score 646.2; DB 12; Length 737;
Best Local Similarity 97.1%; Pred. No. 6, 1e-128;
Matches 702; Conservative 0; Mismatches 13; Indels 8; Gaps 4;

QY 20 GCCGACGTGAGAGAGCGGAGAGAACCTGCGCTCTTTCTTCGACTCCGGAGCGCAGGC 79
    |||||
Db 1 GCCGACGTGAGAGAGCGGAGAGAACCTGCGCTCTTTCTTCGACTCCGGAGCGCAGGC 60
QY 80 TCCAGTGGGAGGCGGACCAATGTTCTCTCAAGAAAGTGGAAACCGGTAGCATGTGG 139
    |||||
Db 61 TCCAGTGGGAGGCGGACCAAGATGTTCTCTCAAGAAAGTGGAAACCGGTAGCATGTGG 120
QY 140 AGCTGGAGCGTGAATGCGATCCTGTGCAATCTGCAGAGGTCAGGTATGGAATGCCCTGC 199
    |||||
Db 121 AGCTGGAGCGTGAATGCGATCCTGTGCAATCTGCAGAGGTCAGGTATGGAATGCCCTGC 180
QY 200 CTTGATGTCAGCTGAAACCAAGCAAGAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 259
    |||||
Db 181 CTTGATGTCAGCTGAAACCAAGCAAGAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 260 CATTCCTTCCACAACTGCTGCATGCTCCCTGTGGTGAACCAACAAATCGCTGCCCTCTG 319
    |||||
Db 241 CATTCCTTCCACAACTGCTGCATGCTCCCTGTGGTGAACCAACAAATCGCTGCCCTCTG 300
QY 320 TGCAGACGAACTGGGTAGTCCAAAGAAATGCGCAATGAGAGGTGGCCGAGGCTGCTG 379
    |||||
Db 301 TGCAGACGAACTGGGTAGTCCAAAGAAATGCGCAATGAGAGGTGGCCGAGGCTGCTG 360
QY 380 GTGTGTTGCTGACCTGCGGAGCAAGACTTAAACACTGCGAGGAGTTCATCTTTGAGAGAGA 439
    |||||
Db 361 GTGTGTTGCTGACCTGCGGAGCAAGACTTAAACACTGCGAGGAGTTCATCTTTGAGAGAGA 420
QY 440 GAGATGCTGTGTCGCGCTTTGAGACTCACCAGAAAGCTTCTTATTAATTTGCTGTGTAG 499
    |||||
Db 421 GAGATGCTGTGTCGCGCTTTGAGACTCACCAGAAAGCTTCTTATTAATTTGCTGTGTAG 480
QY 500 TTTTGGAAATCTCTTACAAATTAAGATTAATTTGTT-AAAAATGCGCTTCTTACCTCTGG 558
    |||||
Db 481 TTTTGGAAATCTCTTACAAATTAAGATTAATTTGTTAAAAATGCGCTTCTTACCTCTGG 540
QY 559 TGTGTGTGTGTGATACGATGATGAAGAGGAGCAACCAAGAAATGATCTTTGTTTA 618
    |||||
Db 541 TGTGTGTGTGTGATACGATGATGAAGAGGAGCAACCAAGAAATGATCTTTGTTTA 600
QY 619 TCTGTACCCAGACT---GGAACATTGTTGTTACAGAAAGACATTGTTGTTATGCT 675
    |||||
Db 601 TCTGTACCCAGACTGTTGGGAAACATATGTTTCAAGAAAGACATG---AGGATATGCT 657
QY 676 TGAGGCTTAAAAATAGATTAACGAATGTTACAGTAAACAATTAATTA-TGCATTGAAAAAGC 734
    |||||
Db 658 TGAGGCTTAAAAATAGATTAACGAATGTTACAGTAAACAATTAATTAATGCAATTGAAAAAGC 717
QY 735 CGA 737
    |||||
Db 718 CGA 720

RESULT 35
BY737338 642 bp mRNA linear EST 17-DEC-2002
LOCUS BY737338 RIKEN full-length enriched, 14 days pregnant adult female
DEFINITION
B1556918
ACCESSION
BY737338.1 GI:27150465
VERSION
BY737338
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 642)
Osakaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaoka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
```


CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM0314 row: o column: 06
High quality sequence stop: 644.

FEATURES

source

1..920
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30243509"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_164"
/note="Vector: pCMV-Sport6.1.ccd; Site 1: EcoRV; Site 2:
NotI; Non-normalized full-length enriched library from
pooled mouse embryonic limb, maxilla and mandible, day
10.5 and 11.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dT. CDNA
enrichment: >1k bp. Average insert size 1.8k bp. Priming
sequence: 5'-GACTAGTTCTAGATCGGAGCGGCCGCTT-3'. Tissue
contributed by, David Rowe. Library constructed by Resgen,
Invitrogen Corp."

BASE COUNT 184 a 152 c 200 g 194 t 190 others

ORIGIN

Query Match 55.0%; Score 627.2; DB 14; Length 920;
Best Local Similarity 98.9%; Pred. No. 7.1e-124;
Matches 642; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 48 GGGCTCTTCTTCGCACTCCGGAGGCGGCAAGCTCGGAGCGCAAGATTTCT 107
DB 1 GGGCTCTTCTTCGCACTCCGGAGGCGGCAAGCTCGGAGCGCAAGATTTCT 60
QY 108 CTCTCAAGAAGTGAACGCGGTAGCATGTGAGCTGGAGCGTTGAGTGCATCTGTG 167
DB 61 CTCTCAAGAAGTGAACGCGGTAGCATGTGAGCTGGAGCGTTGAGTGCATCTGTG 120
QY 168 CCATCTGCAGGCTCAGGTGATGATGCTGCTTGCATGTCAGCTGAAAACAGCAAG 227
DB 121 CCATCTGCAGGCTCAGGTGATGATGCTGCTTGCATGTCAGCTGAAAACAGCAAG 180
QY 228 AGGACGTGTGTGTGCTGGGAGAGTGAACATCTTCCCACTGCTGATGCTCC 287
DB 181 AGGACGTGTGTGTGCTGGGAGAGTGAACATCTTCCCACTGCTGATGCTCC 240
QY 288 TGTGGGTGAACAGAACATCGCTGCTGTGCGCAGCAGCACTGGTATGTCANAAGA 347
DB 241 TGTGGGTGAACAGAACATCGCTGCTGTGCGCAGCAGCACTGGTATGTCANAAGA 300
QY 348 TCGGCAATGAGAGTGGCCGACGCTCTGTGTGTGTTGCTGACCTGGAACAAGACT 407
DB 301 TCGGCAATGAGAGTGGCCGACGCTCTGTGTGTGTTGCTGACCTGGAACAAGACT 360
QY 408 AAAACATGAGAGGATTCATCTTGAAGAGAGAGATGCTGCGCTTTGAGATCAG 467
DB 361 AAAACATGAGAGGATTCATCTTGAAGAGAGAGATGCTGCGCTTTGAGATCAG 420
QY 468 CAAGGCTGCTTATTAATTTGTCTGTAGTTTGGGAAATTTCTCAATTAAGATA 527
DB 421 CAAGGCTGCTTATTAATTTGTCTGTAGTTTGGGAAATTTCTCAATTAAGATA 480
QY 528 ATTGTTAAATATGCTTTCTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 587
DB 481 ATTGTTAAATATGCTTTCTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 588 AGGAGAACACGAGAAATGATTTGTGTATTTGTAACCAAGACT---GGAACATTGG 644
DB 541 AGGAGAACACGAGAAATGATTTGTGTATTTGTAACCAAGACTGTGGGAAACATTGG 600
QY 645 TTCACAGAAACATTTGTTGTATGCTTGAAGGTTAAAAATAGA 693

DB 601 TTCACAGAAACATTTGTTGTATGCTTGAAGGTTAAAAATAGA 649

RESULT 37
BF302267 822 bp mRNA linear EST 21-NOV-2000
LOCUS 602031233f1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4166288 5',
DEFINITION mRNA sequence.
ACCESSION BF302267.1 GI:11248790
VERSION BF302267
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE 1 (bases 1 to 822)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM9455 row: a column: 09
High quality sequence stop: 738.

FEATURES

source

1..822
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4166288"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1lb="NCI_CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-Sport6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 227 a 178 c 229 g 188 t

ORIGIN

Query Match 54.8%; Score 624.4; DB 10; Length 822;
Best Local Similarity 96.9%; Pred. No. 2.8e-123;
Matches 713; Conservative 0; Mismatches 11; Indels 12; Gaps 7;

QY 6 GGGCGCGCGCATGCGCCGACGTGAGAGCGGAGGAACCTGCTCTTCTTGCACT 65
DB 5 GGGCGCGCGCATGCGCCGACGTGAGAGCGGAGGAACCTGCTCTTCTTGCACT 64
QY 66 CCGGAGCGCGAGGCTCCAGTGGGAGGAGGAGAAATGTTCTCTCAAGAAGTGAAG 125
DB 65 CCGGAGCGCGAGGCTCCAGTGGGAGGAGGAGAAATGTTCTCTCAAGAAGTGAAG 124
QY 126 CGGTAGCCATGTGAGCTGGGAGCTTGAAGTGCATCTGCAATGCAAGGTCAGG 185
DB 125 CGGTAGCCATGTGAGCTGGGAGCTTGAAGTGCATCTGCAATGCAAGGTCAGG 184
QY 186 TGATGATGCTGCTGCTTGCATGTCAAGCTGAAAACAAGCAGAGAGCTGTGTGTCT 245
DB 185 TGATGATGCTGCTGCTTGCATGTCAAGCTGAAAACAAGCAGAGAGCTGTGTGTCT 244
QY 246 GGGGAGAGTGAACATCTTCCCACTGTCATGTCATGTCCTGTGGGAGTGAACAAGA 305
DB 245 GGGGAGAGTGAACATCTTCCCACTGTCATGTCATGTCCTGTGGGAGTGAACAAGA 304
QY 306 ATGCTGCTCTGTGCGCAGCAGCACTGGGTATGCCAAGAATCGGCAATGAGAGTGG 365

182 a 167 c 218 g 188 t

ACCESSION	BEG28111	
VERSION	BEG28111.1	GI:9908559
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;	
REFERENCE	1 (bases 1 to 620)	

Query Match	52.3%	Score 596	DB 9	Length 651
Best Local Similarity	99.1%	Pred. No. 3,48-117		
Matches 642	Conservative 0	Mismatches 0	Indels 6	Gaps 4
QY	495	TTTAGTTTGGGAATTCCTCTACATTAAGAATAATTTGTTAAAAATAGCCTTCCATACCT	554	
Db	651	TTTAGTTTGGGAATTCCTCTACATTAAGAATAATTTGTTAAAAATAGCCTTCCATACCT	592	
QY	555	CTGGT	614	
Db	591	CTGGT	532	
QY	615	TTTATCTGTACCCAGCACT--GGACATGTGTTCACAGAGAACATTGTTGTGTGTTA	671	
Db	531	TTTATCTGTACCCAGCACTGTGGAGCAATTGTGTGTTCACAGAGAACATTGTTGTGTTA	472	
QY	672	TGCTTTGAGGGTTAAAAATAGATTAAGAAAGCTTACAGTAAACAAATTAATAGCATTA	731	
Db	471	TGCTTTGAGGGTTAAAAATAGATTAAGAAAGCTTACAGTAAACAAATTAATAGCATTA	412	
QY	732	AGCCGACTCCTCCTAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	791	
Db	411	AGCCGACTCCTCCTAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	352	
QY	792	CTTCATTTGCTGT	851	
Db	351	CTTCATTTGCTGT	292	
QY	852	TAAACTGTAAATATGCGCTATCTGTGCGGTAAACGGCTTTGTCTCTGACTTCTCCATCT	911	
Db	291	TAAACTGTAAATATGCGCTATCTGTGCGGTAAACGGCTTTGTCTCTGACTTCTCCATCT	232	
QY	912	TGACTTTGGCCAGAGAGCTGTGATTTGTCAACCACTTACTTGTCTTAAAGAACTGTCTGT	971	
Db	231	TGACTTTGGCCAGAGAGCTGTGATTTGTCAACCACTTACTTGTCTTAAAGAACTGTCTGT	172	
QY	972	TTTGCCGAAAGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG	1031	
Db	171	TTTG--CGAAGGT	115	
QY	1032	GTATTAACACTGAAGTTCATTATGCAATGTTTAAATAAATATGTCCTTTGATTTAA	1091	
Db	114	GTATTAACACTGAAGTTCATTATGCAATGTTTAAATAAATATGTCCTTTGATTTAA	55	
QY	1092	AGTTTGATTTATATCTTAAATCAATTAACATTAATTCATCAATTAAT	1139	
Db	54	AGTTTGATTTATATCTTAAATCAATTAACATTAATTCATCAATTAAT	7	
RESULT 41				
LOCUS	CB951114	728 bp	mRNA	linear
DEFINITION	AGENCOUNT_13442509 NIH_MGC_177 Mus musculus cDNA clone			
ACCESSION	CB951114			
VERSION	CB951114.1	GI:30206005		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 728)			
AUTHORS	NIH-MGC http://imgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-r@mail.nih.gov Tissue Procurement: Dr. Michael Brownstein cDNA Library Preparation: The Michael Brownstein Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:			

FEATURES	Source
http://image.jini.gov Plate: NDCM06 row: 1 column: 15 High quality sequence stop: 531. Location/Qualifiers	1. .728 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /clone="IMAGE:30316022" /lab_host="DH10B (rt1-phage-resistant)" /clone_id="NIH MGc.177" /note="Organ: liver; Vector: PDNR-LIB; Site: 1; SiteII (ggccatgatggcc); Site: 2; SiteI (ggccgctctggcc); cDNA made by oligo-dt priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGGTATCAACGACAGATGGCATTCAGCCCGG-3' and 5'-ATTCTAGAGCCGAGGGGGCCGACATG-dt(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5 kb size fraction. Library created in the laboratory of M. Brownstein (NIMH, NIH). Note: this is a NIH MGc Library."
BASE COUNT	175 a 157 c 215 g 181 t
ORIGIN	
Query Match	52.0%; Score 592.4; DB 14; Length 728;
Best Local Similarity	98.1%; Freq. No. 2e-116;
Matches 610; Conservative 0; Mismatches 11; Indels 1; Gaps 1	
QY	1 GTTTCGCCCGCCCGCCGATGCGCGAGAGACGCGAGAAACCTTCGTCCTTTCTTC 60
Db	9 GTTTCGCCCGCCCGCCGATGCGCGAGAGACGCGAGAAACCTTCGTCCTTTCTTC 68
QY	61 GCATTCGGGGAGCGCAGCTCCAACTCGGGAGCGCAACAAGTGTCTCTCCAAAGAATG 120
Db	69 GCATTCGGGGAGCGCAGCTCCAACTCGGGAGCGCAACAAGTGTCTCTCCAAAGAATG 128
QY	121 GAACGCGGTAAGCATGTGAGCTGGAGCGTTGAGTGCATCTTGCCATCTGCAGGCT 180
Db	129 GAACGCGGTAAGCATGTGAGCTGGAGCGTTGAGTGCATCTTGCCATCTGCAGGCT 188
QY	181 CCAGGTGATGATGTCCTGCTGCATGTCCAACTGAAAACAAGCAAGAGACTGTGTCT 240
Db	189 CCAGGTGATGATGTCCTGCTGCATGTCCAACTGAAAACAAGCAAGAGACTGTGTCT 248
QY	241 GGTCTGGGGAGAGGTAAACCATCTTCCACAAACTGTCATGTCCCTGTGGGTGAACA 300
Db	249 GGTCTGGGGAGAGGTAAACCATCTTCCACAAACTGTCATGTCCCTGTGGGTGAACA 308
QY	301 GAACATCGCTCCCTCTGTGTCACAGACAGACTGGTAGTCCAAAGATGGCAATGAA 360
Db	309 GAACATCGCTCCCTCTGTGTCACAGACAGACTGGTAGTCCAAAGATGGCAATGAA 368
QY	361 GGTGGCCACAGCGCTCTGTGTGTTGCTGTACCTCTGACCAAGACTAAACCTGAGGG 420
Db	369 GGTGGCCACAGCGCTCTGTGTGTTGCTGTACCTCTGACCAAGACTAAACCTGAGGG 428
QY	421 GATTCTCTTGAAGAGAGAGAGATCTGTGCGCTTTGAGACTCAACAAAGCTTGCTT 480
Db	429 GATTCTCTTGAAGAGAGAGATCTGTGCGCTTTGAGACTCAACAAAGCTTGCTT 488
QY	481 TATTAAATTTGCTGTTAGTTTGGAAATTTCTTCAATTAATTAATTAATTTGTT-AAAA 539
Db	489 TATTAAATTTGCTGTTAGTTTGGAAATTTGCTTCAATTAATTAATTAATTTGTTAAAA 548
QY	540 TGGCCTTTCTTACCTCTGTGTGTGTGTGATACGAATGATAGAAAGAGCGAAGACAC 599
Db	549 TGGCCTTTCTTACCTCTGTGTGTGTGTGATACGAATGATAGAAAGAGCGAAGACAC 608
QY	600 AGAAATGATCTTTGTTATCT 621
Db	609 AGAAATGATCTTTGTTATCT 630

	EST	42
B757681	BY57681	653 bp mRNA linear EST 17-DEC-2002
LOCUS	BY57681 RIKEN full-length enriched, 14 days pregnant adult female placenta Mus musculus cDNA clone I53002SD07 3', mRNA sequence.	
DEFINITION		
ACCESSION	BY57681	
VERSION	BY57681.1 GI:27191292	
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.	
REFERENCE	Oikarinen, Y., Furumori, M., Kasugawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Osato, N., Saito, R., Suzuki, H., Yamanaoka, I., Miyosawa, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sullivan, R., Takenaki, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Vercardo, R., Wagner, L., Wahlerstedt, C., Wang, Y., Watanebe, Y., Wells, C., Wilting, L. G., Wynshaw-Borris, A., Yanagisawa, M., Yang, F., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Caminici, P., Hayatsu, N., Hirozane-Kobayashi, T., Konno, H., Nakamura, M., Sakakura, N., Sato, K., Shiraki, T., Wakai, K., Kawaji, T., Akizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sekai, K., Sasaki, D., Shibata, E. S., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs <i>Nature</i> 420, 563-573 (2002)	
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Adachi, J., Akizawa, K., Akimura, T., Arakawa, T., Carinci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirazane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawasaki, J., Kohjima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Nunawake, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Saeki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeuchi, Y., Wakai, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences <i>Mamm. Genome</i> . 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. <i>Genome Res.</i> 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. <i>Genome Res.</i> 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. <i>Genome Res.</i> 11 (2), 281-289 (2001) cDNA Library was prepared and sequenced in Mouse Genome	

		Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome-gsc.riken.go.jp) for further details.	
FEATURES	Location/Qualifiers		
SOURCE	1. 653		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="IS30025D07"		
	/sex="Female"		
	/tissue_type="placenta"		
	/dev_stage="14 days pregnant adult"		
	/clone_lib="RIKEN full-length enriched, 14 days pregnant adult Female placenta"		
BASE COUNT	191 a 103 c 135 g 223 t	1 others	
ORIGIN			
Query Match	51.5%; Score 586.8; DB 14; Length 653;		
Best Local Similarity	98.5%; Pred. No. 3,2e-115;		
Matches	645; Conservative 0; Mismatches 3; Indels 7; Gaps 5;		
QY	480 TTATTAATTTGTCGTGTAGTTAGTTTGGGAAATTCCTCAATTAAGATTAATTTGTTAAAAA	539	
Db	3 TTATTAATTTGTCGTGTAGTTAGTTTGGGAAATTCCTCAATTAAGATTA- TTGTTAAAAA	61	
QY	540 TGGCCTTTCCACCTCGTGCTGTGTGTGTGATGCGAATGATGAAGAGGACAAACC	599	
Db	62 TGGCCTTTCCACCTCGTGCTGTGTGTGTGATGCGAATGATGAAGAGGACAAACC	121	
QY	600 AGAAAAATGATTTGTTATCTGTACCAACGACAT---GGAACATTTGTTACAGAGAA	656	
Db	122 AGAAAAATGATTTGTTATCTGTACCAACGACATGTGGAAACATTTGTTACAGAGAA	181	
QY	657 CATTGTTTGTGTTTATGCTTGAGGGTTAAAAAATAGATTAACGATGTTACGTAAACAA	716	
Db	182 CATTGTTTGTGTTTATGCTTGAGGGTTAAAAAATAGATTAACGATGTTACGTAAACAA	241	
QY	717 TAAAAATGATTGAAAAGCCGACCTCTCTTAATCCTTTTGTGTGGAGAGAGGCAACG	776	
Db	242 TAAAAATGATTGAAAAGCCGACCTCTCTTAATCCTTTTGTGTGGAGAGGCAACG	301	
QY	777 AGGCACACCTGCTGCTTCAATTCCTGTGATGAGATTTTAACCTGACCTCAGTGAAGA	836	
Db	302 AGGCACACCTGCTGCTTCAATTCCTGTGATGAGATTTTAACCTGACCTCAGTGAAGA	361	
QY	837 GGCCTAACTGTGCGGTAACTGTATATATGCGTAACCTGTGCGGTAAACGCGTTTGTCTCC	896	
Db	362 GGCCTAACTGTGCGGTAACTGTATATATGCGTAACCTGTGCGGTAAACGCGTTTGTCTCC	421	
QY	897 TGACTCTTCATCTTTTGACTTGGCCAGAAACCTCGATGTTTCAACCACTTAAGTTCTTAA	956	
Db	422 TGACTCTTCATCTTTTGACTTGGCCAGAAACCTCGATGTTTCAACCACTTAAGTTCTTAA	481	
QY	957 GAACGTGTTTTGTGTTTTGCGGAGGTGTATGTATGTTTAACTCAAAAAATATTAGAG	1016	
Db	482 GAACGTGTTTTGTGTTTTGCGGAGGTGTATGTATGTTTAACTCAAAAAATATTAGAG	539	
QY	1017 GAAATGAGCTTACTAGTATAACACTGGAAGTTCATTATGCAATGTTTTTAATTAATATTGT	1076	
Db	540 GAAATGAGCTTACTAGTATAACACTGGAAGTTCATTATGCAATGTTTTTAATTAATATTGT	598	
QY	1077 GCTTTGAGTTATTTAAAGTTTGATATATATCTTTAAATCATTAAACTTAATTCATC	1131	
Db	599 GCTTTGAGTTATTTAAAGTTTGATATATATCTTTAAATCATTAAACTTAATTCATC	653	

DEFINITION UT-M-EMO-bv1-p-18-0-UI.r1 NIH_BMAP_EMO Mus musculus cDNA clone
IMAGE:5692673 5', mRNA sequence.
ACCESSION BM946344
VERSION BM946344.1 GI:19429929
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
source
Seq primer: pyx-5.
Location/Qualifiers
1..624
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5692673"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_1lb="NIH_BMAP_EMO"
/note="Organ: Brain; Vector: pyx-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCCACGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System, supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 150 a 134 c 180 g 158 t 2 others
ORIGIN

Query Match 51.2%; Score 584; DB 12; Length 624;
Best Local Similarity 99.5%; Pred. No. 1.3e-114;
Matches 606; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

27 TGAAGAGCGGCGAGGACCTGCGCTCTTCTTGCACTC-CGGAGCGCAGGCTCCAG 85
Db 1 TGAAGAGCGGCGAGGAA-CTGCGCTCTTCTTGCACTCNCGGAGCGCGCTCCAG 59
Qy 86 TCGGAGGCGACAAATGTTCTCTCAAGAGTGGAA-CGGGTAGGCATGTGAGCTGG 145
Db 60 TCGGAGGCGACAAATGTTCTCTCAAGAGTGGAA-CGGGTAGGCATGTGAGCTGG 119
Qy 146 GACGTTGAGTGCATCTGTCATCGACGGGTCCAGGTGATGATGATGCTGCTTGA 205
Db 120 GACGTTGAGTGCATCTGTCATCGACGGGTCCAGGTGATGATGATGCTGCTTGA 179
Qy 206 TGTCAAGTGAAGAACAGAGAGAGAGCTGTGTGTGTGTGGGAGAGATGAACATCC 265
Db 180 TGTCAAGTGAAGAACAGAGAGAGAGCTGTGTGTGTGTGGGAGAGATGAACATCC 239

Qy 266 TTCCACAACTGCTGCATGTCCTGTGGTGAAACGAAACAATCGTCCCTGTGCCAG 325
Db 240 TTCCACAACTGCTGCATGTCCTGTGGTGAAACGAAACAATCGTCCCTGTGCCAG 299
Qy 326 CAGACTGGGTGTGTCGAAAGAAATCGGCAAAATAGAGAGTGGCCAGCGCTCTGTGTGG 385
Db 300 CAGACTGGGTGTGTCGAAAGAAATCGGCAAAATAGAGAGTGGCCAGCGCTCTGTGTGG 359
Qy 386 TTGCTGACCTGGACAAAGACTAAACATCGACGGGATTCATCTTTGAGAGAGAGAT 445
Db 360 TTGCTGACCTGGACAAAGACTAAACATCGACGGGATTCATCTTTGAGAGAGAGAT 419
Qy 446 GCTGTGGCTTTGAGACTCAACCAAGGCTGCTTATTAATTTGCTGTAGTTTGG 505
Db 420 GCTGTGGCTTTGAGACTCAACCAAGGCTGCTTATTAATTTGCTGTAGTTTGG 479
Qy 506 GAAATTCCTACATTAAGTAAATTTGTTAAATAATGGCTTTCTTACTCTGTGTGTGT 565
Db 480 GAAATTCCTACATTAAGTAAATTTGTTAAATAATGGCTTTCTTACTCTGTGTGTGT 539
Qy 566 GTGTGATGCAATGATAGAGAGCGGAGAACACCAAAATGATTTGTTATCTGTAC 625
Db 540 GTGTGATGCAATGATAGAGAGCGGAGAACACCAAAATGATTTGTTATCTGTAC 599
Qy 626 CCACGACTG 634
Db 600 CCACGACTG 608

RESULT 44
CB318426 757 bp mRNA linear EST 04-MAR-2003
LOCUS CB318426
DEFINITION AGNCOURT 12396411 NIH MGC 165 Mus musculus cDNA clone
IMAGE:30283955 5', mRNA sequence.
ACCESSION CB318426
VERSION CB318426.1 GI:2842661
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. Leslie L. Heckert
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: NDCM47 row: d column: 12
High quality sequence stop: 500.
Location/Qualifiers
1..757
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30283955"
/tissue_type="Primary cultures of Sertoli cells"
/lab_host="DH10B (TI-phage-resistant)"
/clone_1lb="NIH MGC 165"
/note="Organ: testis; Vector: pDNR-LIB; Site 1: SfiI (ggccatratggcc); Site 2: SfiI (ggcgccctggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCGCATATGAGC-3' and 3' adaptor sequence: 5'-ATTGAGAGCGGCGGCGGCGCATATG-dT(30)BN-3' (where B = A, C, G or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.6-3.5 kb). 15/15 colonies contained inserts by

BASE COUNT 190 a 169 c 213 g 204 t
ORIGIN

Query Match 50.2%; Score 572.8; DB 14; Length 776;
Best Local Similarity 98.5%; Pred. No. 3.2e-112;
Matches 599; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

```

QY 14 GCCATGCGGACGCTGAGAGAGCGGAGGAAACCTGCGCTCTTCTCGACCTCGGAGAC 73
DB 4 GCCATGCGGACGCTGAGAGAGCGGAGGAAACCTGCGCTCTTCTCGACCTCGGAGAC 63
QY 74 GCAAGCTCCAAAGTCGGAGGCGCAAGATGTTCTCTCAAGAGTGAACGGCGTAGCC 133
DB 64 GCAAGCTCCAAAGTCGGAGGCGCAAGATGTTCTCTCAAGAGTGAACGGCGTAGCC 123
QY 134 ATGTGAGCTGGAGAGCTTGAAGTCCGATCTGTCGCACTGCGAGGTCGAGTGAAT 193
DB 124 ATGTGAGCTGGAGAGCTTGAAGTCCGATCTGTCGCACTGCGAGGTCGAGTGAAT 183
QY 194 GCGTGCCTTCGATGTCAGCTGAAGTGAACCAAGCAAGAGAGAGCTGTGTGTGGTGGGAGAG 253
DB 184 GCGTGCCTTCGATGTCAGCTGAAGTGAACCAAGCAAGAGAGAGCTGTGTGTGGTGGGAGAG 243
QY 254 TGTAAACATTCCTTCCAAACAGCTGCTGATGTCCTGTGGGTGAACAGAACATCGCTGC 313
DB 244 TGTAAACATTCCTTCCAAACAGCTGCTGATGTCCTGTGGGTGAACAGAACATCGCTGC 303
QY 314 CCTGTGTCCAGAGAGAGCTGGTGAATGCCAAAGATGAGAGTGGCGCCGAGGCG 373
DB 304 CCTGTGTCCAGAGAGAGCTGGTGAATGCCAAAGATGAGAGTGGCGCCGAGGCG 363
QY 374 CTCCTGTGTGTGTGTGTGTCGACCTTGAAGCAAGACTAACAAGTGAAGTCACTCTGA 433
DB 364 CTCCTGTGTGTGTGTGTGTCGACCTTGAAGCAAGACTAACAAGTGAAGTCACTCTGA 423
QY 434 GAGAGAGAGAGAGCTGTGCGCCCTTTGAGACTCAACCAAGGCTTGTATTAATTTGCT 493
DB 424 GAGAGAGAGAGAGCTGTGCGCCCTTTGAGACTCAACCAAGGCTTGTATTAATTTGCT 483
QY 494 GTTTAGTTTGGGAATTTCTTCAATTAAAGTAATTTG-TTAAATGGCCTTCTTAC 552
DB 484 GTTTAGTTTGGGAATTTCTTCAATTAAAGTAATTTGTTTAAATGGCCTTCTTAC 543
QY 553 CTCGTGTGTGTGTGTGTGATGATGCAATGACATA-GAAGAGGAGAGACCGCAAAATGATCT 611
DB 544 CTCGTGTGTGTGTGTGTGATGATGCAATGACATA-GAAGAGGAGAGACCGCAAAATGATCT 603
QY 612 TTGTTTAT 619
DB 604 TTTGTTTT 611

```

RESULT 48
BI153818 735 bp mRNA linear EST 05-JUL-2001
LOCUS 602870803F1 NCI_GCAP_Mam2 Mus musculus cDNA clone IMAGE:5002729 5',
DEFINITION
BI153818 mRNA sequence.
ACCESSION
BI153818 GI:14613819
VERSION
BI153818.1
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 735)
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1038 row: e column: 02
High quality sequence stop: 622.
Location/Qualifiers

FEATURES
source 1..735
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5002729"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_11b="NCI_GCAP Mam2"
/note="Organ: mammary; Vector: pCMV-Sport6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 173 a 185 c 201 g 176 t
ORIGIN

Query Match 49.9%; Score 568.8; DB 12; Length 735;
Best Local Similarity 96.9%; Pred. No. 2.3e-111;
Matches 624; Conservative 0; Mismatches 12; Indels 8; Gaps 4;

```

QY 20 GCCAGCTGAGAGAGCGGAGGAAACCTGCGCTTCTTCTCGACTCCGGAGCGGAGGC 79
DB 1 GCCAGCTGAGAGAGCGGAGGAAACCTGCGCTTCTTCTCGACTCCGGAGCGGAGGC 60
QY 80 TCCAA-GTGCGGAGGCGCAACATGTTCTCTTCAAGAGTGAACCGCGTACCATGTG 138
DB 61 TCCAAATGTCGGAGGCGCAACATGTTCTCTTCAAGAGTGAACCGCGTACCATGTG 120
QY 139 GAGCTGGAGAGCTTGAAGTGAATCCTGTGCCATCTGCAAGGTCGAGGTGATGATGCTG 198
DB 121 GAGCTGGAGAGCTTGAAGTGAATCCTGTGCCATCTGCAAGGTCGAGGTGATGATGCTG 180
QY 199 CCTTGATGTCAGGCTGAACCAAGCAAGAGAGCTGTGTGTGTGTGTGGTGGGAGAGTGA 258
DB 181 CCTTGATGTCAGGCTGAACCAAGCAAGAGAGCTGTGTGTGTGTGTGGTGGGAGAGTGA 240
QY 255 CCATTCCTTCCAAACTGTGCTATGCTCTGTGGGTGAACCAAGATCGCTGCTCT 318
DB 241 CCATTCCTTCCAAACTGTGCTATGCTCTGTGGGTGAACCAAGATCGCTGCTCT 300
QY 319 GTGCCAGAGAGCTGGGTAGTCCAAAGATGGGCAATGAGAGGTGGCCAGGCGCTCT 378
DB 301 GTGCCAGAGAGCTGGGTAGTCCAAAGATGGGCAATGAGAGGTGGCCAGGCGCTCT 360
QY 379 GGTGTGTGTGTGTGTGATGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
DB 361 GGTGTGTGTGTGTGTGATGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 437 AGAGAGAGAGCTGTGCGCCTTTGAGACTCACCAGAGGCTTGTATTAATTTGCTGT 496
DB 421 AGAGAGAGAGCTGTGCGCCTTTGAGACTCACCAGAGGCTTGTATTAATTTGCTGT 480
QY 497 TAGTTTGGGAATTTCTTCAATTAAGATA-ATTGTTAAATAGGCTTCTTCAATCCT 555
DB 481 TAGTTTGGGAATTTCTTCAATTAAGATA-ATTGTTAAATAGGCTTCTTCAATCCT 540
QY 556 TGTGTGTGTGTGTGTGATGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615
DB 541 TGTGTGTGTGTGTGTGATGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 616 TTATCTGTAACCAAGACT---GGAACATGTGTTCAAGAGA 655
DB 601 TTATCTGTAACCAAGACTGTGGAGACATTTGTCTACAGAGA 644

```

[illegible]

QY	344	AGAAATCGGCAATAGAGATGTCGCCAGAGCGCTCCTGTGTGTGTGTCTGACACCTGGACAAA	403
Db	301	AGAAATCGGCAATAGAGATGTCGCCAGAGCGCTCCTGTGTGTGTGTCTGACACCTGGACAAA	360
QY	404	GACTTAACACTGCGACGGGATTCATCTCTTGAGAGAGAGAGATGCTGCGCCCTTTGAGAC	463
Db	361	GACTTAACACTGCGACGGGATTCATCTCTTGAGAGAGAGAGATGCTGCGCCCTTTGAGAC	420
QY	464	TCACCAAGAGGCTGCTTTATTTATTTGTCTGTTTATGTTTGGGAAATCTCTCAATTTAA	523
Db	421	TCACCAAGAGGCTGCTTTATTTATTTGTCTGTTTATGTTTGGGAAATCTCTCAATTTAA	480
QY	524	GATTAATTTGTTAAATATGCGCTTCTCTACCTGCTGCTGTGTGTGTGTGATACGAATGATA	583
Db	481	GATTAATTTGTTAAATATGCGCTTCTCTACCTGCTGCTGTGTGTGTGTGATACGAATGATA	540
QY	584	GATTAATTTGTTAAATATGCGCTTCTCTACCTGCTGCTGTGTGTGTGTGATACGAATGATA	641
Db	541	NAAGACGAGAAACACCAAGAAATGATCT	568
RESULT	50		
LOCUS	AA230335	577 bp	mRNA linear EST 26-FEB-1997
DEFINITION	nm04c09.r1 Soares mouse 3NME12.5 Mus musculus cDNA clone		
ACCESSION	AA230335		
VERSION	AA230335.1	GI:1852740	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B., Thelshing, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.		
JOURNAL	The WashU-HMNI Mouse EST Project		
COMMENT	Unpublished Contract: Marra M/Mouse EST Project WashU-HMNI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNC, contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:409416 Seq primer: -28m3 rev2 ET from Amersham High quality sequence scop: 498. Location/Qualifiers 1. 577 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:669712" /sex="unknown" /tissue_type="fetus" /dev_stage="12.5dpc total fetus" /lab_host="DH10B" /clone_1ib="Soares mouse 3NME12.5" /note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI 1st strand cDNA was primed with a Not I - oligo(dT) primer [5], TgTtACCAATCTGAAGTGGAGCGCGCTTATTTTATTTTATTTT 3'1, on total mouse RNA [provided by Minoru Ko, Wayne State Univ.]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into		

the Not I and Eco RI sites of the modified p773 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT

142 a 124 c 164 g 147 t

ORIGIN

Query Match

48.8%; Score 556.8; DB 9; Length 577;

Best Local Similarity 99.5%; Pred. No. 8.5e-109;

Matches 569; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
QY 63 ACTCCGGAGCGCCAGGCTCCAGTCCGGAGGCGACAAAGATGTTCTCTTCAGAGATGGA 122
DB 6 ACTCCGGAGCGCGAGGCT-CAAGTCGGAGGCGACAAAGATGTTCTCTTCAGAGAGTGA 64
QY 123 ACCGGTAGCCATGTGGAGCTGGAGCTGAGTGCATACCTGTCATCTGCAGGGTCC 182
DB 65 ACCGGTAGCCATGTGGAGCTGGAGCTGAGTGCATACCTGTCATCTGCAGGGTCC 124
QY 183 AGGTATGATGCTGCTGCTTCGATGTCAGCTGAAAAACAAGACAGACTGTGTGTGG 242
DB 125 AGGTATGATGCTGCTGCTTCGATGTCAGCTGAAAAACAAGACAGACTGTGTGTGG 184
QY 243 TCTGGGAGAGTGAACCATTCCTTCCACACTGCTGCATGCTCTGTGGTGAACAGA 302
DB 185 TCTGGGAGAGTGAACCATTCCTTCCACACTGCTGCATGCTCTGTGGTGAACAGA 244
QY 303 ACAATGCTGCCCTCTGTGCCAGACAGACTGGGTAGTCCAAAGAAATCGGCAATGAGAG 362
DB 245 ACAATGCTGCCCTCTGTGCCAGACAGACTGGGTAGTCCAAAGAAATCGGCAATGAGAG 304
QY 363 TGGCCAGGCGCTCTGTGTGTGTGCTGACCTGGACAAAGACTAAACACTGCAGGGGA 422
DB 305 TGGCCAGGCGCTCTGTGTGTGTGCTGACCTGGACAAAGACTAAACACTGCAGGGGA 364
QY 423 TTGATCTTGAAGAGAGATGCTGTGCGCTTTGAGACTGACCAAGGCTTGCTTGA 482
DB 365 TTGATCTTGAAGAGAGATGCTGTGCGCTTTGAGACTGACCAAGGCTTGCTTGA 424
QY 483 TTAATTGCTGTTAGTTTGGGAAATCTCTACAAATTAAGTAATTTGTTAAAAATGG 542
DB 425 TTAATTGCTGTTAGTTTGGGAAATCTCTACAAATTAAGTAATTTGTTAAAAATGG 484
QY 543 CCTTTCCTACCTCTGTGTGTGTGTGATACGAATGATAGAGAGAGAACACCAGA 602
DB 485 CCTTTCCTACCTCTGTGTGTGTGTGATACGAATGATAGAGAGAGAACACCAGA 544
QY 603 AAATGATCTTTGTTATCTGTACCCAGACTG 634
DB 545 AAATGATCTTTGTTATCTGTACCCAGACTG 576
```

Search completed: November 7, 2003, 15:15:38
Job time : 2889.08 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 11:58:22 ; Search time 478.215 Seconds
(without alignments)
827.407 Million cell updates/sec

Title: US-09-509-779-3_COPY_141_264

Perfect score: 124
Sequence: 1 CGATACGTCGCGCCATCTGCA.....ATTCTTCACAACTGCTGC 124

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 2141354 seqs, 159547879 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Published Applications NA.*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	124	100.0	342	9	US-09-826-312-7	Sequence 7, Appl1
2	124	100.0	342	14	US-10-108-767-7	Sequence 7, Appl1
3	124	100.0	342	14	US-10-152-156-7	Sequence 7, Appl1
4	124	100.0	612	10	US-09-796-692-8547	Sequence 8547, Appl
5	124	100.0	612	14	US-10-040-862-8547	Sequence 8547, Appl
6	124	100.0	836	10	US-09-764-864-39	Sequence 39, Appl
7	124	100.0	836	10	US-09-764-864-498	Sequence 498, Appl
8	95	76.6	535	14	US-10-102-524-396	Sequence 396, Appl
9	50.8	1024	14	US-10-198-846-12407	Sequence 12407, A	
10	35	28.2	60	12	US-09-908-975-13285	Sequence 13285, A
11	24	19.4	933	14	US-10-198-846-7306	Sequence 7306, Appl
12	19	15.3	45	12	US-09-974-026-61	Sequence 61, Appl
13	19	15.3	3372	12	US-10-027-632-113318	Sequence 113318, A
14	19	15.3	3372	13	US-10-027-632-113318	Sequence 113318, A
15	18	14.5	336	11	US-09-918-995-19059	Sequence 19059, Appl
16	18	14.5	541	10	US-09-998-598-450	Sequence 450, Appl

17	18	14.5	930	12	US-10-027-632-31061	Sequence 31061, A
18	18	14.5	930	13	US-10-027-632-31061	Sequence 31061, A
19	18	14.5	2630	14	US-10-106-868-799	Sequence 799, App
20	17	13.7	252	12	US-10-027-632-276727	Sequence 276727, App
21	17	13.7	252	13	US-10-027-632-276727	Sequence 276727, App
22	17	13.7	355	9	US-09-864-761-4974	Sequence 4974, App
23	17	13.7	355	9	US-09-864-761-21703	Sequence 21703, A
24	17	13.7	464	14	US-10-066-543-1656	Sequence 1656, App
25	17	13.7	500	11	US-09-918-995-31855	Sequence 31855, A
26	17	13.7	712	12	US-10-027-632-145248	Sequence 145248, A
27	17	13.7	712	13	US-10-027-632-145248	Sequence 145248, A
28	17	13.7	833	12	US-10-027-632-162982	Sequence 162982, A
29	17	13.7	833	12	US-10-027-632-162982	Sequence 162983, A
30	17	13.7	833	13	US-10-027-632-162982	Sequence 162982, A
31	17	13.7	833	13	US-10-027-632-162983	Sequence 162983, A
32	17	13.7	1781	12	US-10-027-632-255720	Sequence 255720, A
33	17	13.7	1781	13	US-10-027-632-255720	Sequence 255720, A
34	17	13.7	2904	14	US-10-281-904-3	Sequence 3, Appl
35	17	13.7	3494	10	US-09-880-107-2295	Sequence 2295, App
36	17	13.7	3494	14	US-10-205-823-31	Sequence 31, Appl
37	17	13.7	3494	14	US-10-264-374-200	Sequence 200, App
38	17	13.7	3494	14	US-10-281-904-1	Sequence 1, Appl
39	17	13.7	3681	12	US-10-133-013-221	Sequence 221, App
40	17	13.7	3686	10	US-09-981-353-121	Sequence 121, App
41	17	13.7	3686	13	US-10-044-090-218	Sequence 218, App
42	17	13.7	3686	14	US-10-094-817-86	Sequence 86, Appl
43	17	13.7	3728	9	US-09-925-297-338	Sequence 338, App
44	17	13.7	3815	12	US-10-133-013-222	Sequence 222, App
45	17	13.7	11150	11	US-09-764-891-6245	Sequence 6245, App

ALIGNMENTS

```
RESULT 1
US-09-826-312-7
; Sequence 7, Application US/09826312
; Patent No. US20020042083A1
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jiating
; APPLICANT: Shuang, Julie
; APPLICANT: Pray, Todd R.
; TITLE OF INVENTION: UBIQUITIN LIGASE ASSAY
; FILE REFERENCE: A-68613-1/RMS/JUD
; CURRENT APPLICATION NUMBER: US/09/826.312
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 09/542.497
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-826-312-7
```

Query Match	100.0%	Score 124;	DB 9;	Length 342;
Best Local Similarity	100.0%	Pred. No. 6.9e-60;		
Matches 124;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	CGATACGTCGCGCCATCTGCAAGGTCAGGTGATGATGCTCTTCTTAGATGTCGAAGCTCA	60	
DB	141	CGATACGTCGCGCCATCTGCAAGGTCAGGTGATGATGCTCTTCTTAGATGTCGAAGCTCA	200	
QY	61	AAACAAAGAGAGAGCTGCTGCTGCTGCGAGGAAGTATATCAATCTCTTCAACAAGTCA	120	
DB	201	AAACAAAGAGAGAGCTGCTGCTGCTGCGAGGAAGTATATCAATCTCTTCAACAAGTCA	260	
QY	121	CTGC 124		
DB	261	CTGC 264		

```

RESULT 2
US-10-108-767-7
; Sequence 7, Application US/10108767
; Publication No. US20030104474a1
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; TITLE OF INVENTION: ASSAYS FOR IDENTIFYING UBIQUITIN AGENTS AND FOR IDENTIFYING AGENTS
; FILE REFERENCE: A-68613-5/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/108,767
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 09/542,497
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 09/826,312
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 10/091,139
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-767-7

```

```

Query Match 100.0%; Score 124; DB 14; Length 342;
Best Local Similarity 100.0%; Pred. No. 6,9e-60;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTCGCCCATCTTCAGAGGTCAGAGTGATGATCCTGTCTTAGATGTCAAGCTGA 60
Db 141 CGATACGTCGCCCATCTTCAGAGGTCAGAGTGATGATGATCCTGTCTTAGATGTCAAGCTGA 200
QY 61 AAACAAACAAAGAGACTGTGTGTGTGTGTGTGGGGAGATGATATCATTTCTTCCAAACTG 120
Db 201 AAACAAACAAAGAGACTGTGTGTGTGTGTGTGGGGAGATGATATCATTTCTTCCAAACTG 260
QY 121 CTGC 124
Db 261 CTGC 264

US-10-152-156-7
US-10-152-156-7
; Sequence 7, Application US/10152156
; Publication No. US20030108947A1
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; TITLE OF INVENTION: ASSAYS FOR IDENTIFYING UBIOUITIN AGENTS AND FOR IDENTIFYING AGENTS
; FILE REFERENCE: A-68613-6/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/152,156
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 09/542,497
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 09/826,312
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 10/091,174
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 10/091,139
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 10/109,460
; PRIOR FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 10/108,767
; PRIOR FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/291,836

```

```

;
; PRIOR FILING DATE 2001-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.
; SEQ ID NO 7
; LENGTH: 342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-152-156-7

```

Query Match	100.0%;	Score 124;	DB 14;	Length 342;
Best Local Similarity	100.0%;	Pred. No. 6.9e-60;		
Matches 124;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	CGATACGTGCGCATCTGCAGGGCTCAGGTATGATGCCGTCTTATGATGTAAGCTGA	60
Db	CGATACGTGCGCATCTGCAGGGCTCAGGTATGATGCCGTCTTATGATGTAAGCTGA	2000
QY	AAACAAACAGAGGACTGTGTGTGTGTGTGGGGAGATGTATCATCTCTTCCAAACTG	120
Db	AAACAAACAGAGGACTGTGTGTGTGTGTGGGGAGATGTATCATCTCTTCCAAACTG	260
QY	CTGC	124
Db	CTGC	264

RESULT 4
US-09-796-692-8547
; Sequence 8547, Application US/09796692
; Publication No. US20020198362A1

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



```

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7306
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 36, 150, 168, 187, 224, 266, 305, 375, 378, 405, 429, 438,
; LOCATION: 439, 454, 458, 464, 472, 485, 491, 495, 516, 536, 537, 538,
; LOCATION: 544, 545, 548, 550, 558, 559, 563, 565, 567, 570, 578, 581,
; LOCATION: 585, 617, 633, 635, 636, 637, 641, 644, 647, 651, 657
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 672, 674, 680, 690, 692, 694, 695, 699, 704, 705, 707, 709,
; LOCATION: 725, 733, 740, 743, 744, 747, 748, 754, 756, 757, 762, 765,
; LOCATION: 769, 771, 774, 775, 777, 783, 784, 786, 788, 792, 795, 797,
; LOCATION: 798, 801, 802, 805, 809, 812, 813, 814, 815, 817, 818
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 819, 824, 827, 836, 838, 846, 851, 852, 855, 860, 865, 871,
; LOCATION: 876, 882, 893, 901, 903, 904, 905, 908, 909, 910, 913, 915,
; LOCATION: 919, 925, 927, 930, 931
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-7306

Query Match          19.4%; Score 24; DB 14; Length 933;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      101 AATCATTCCTCCACACACTGCTGC 124
        |||||||
DB       90 AATCATTCCTCCACACACTGCTGC 113

RESULT 12
US-09-974-026-61/c
; Sequence 61, Application US/09974026
; Publication No. US20030194398A1
; GENERAL INFORMATION:
; APPLICANT: Tamburini, Paul P
; APPLICANT: Davis, Gary
; APPLICANT: Delaria, Katherine A
; APPLICANT: Christopher, Marlor W
; APPLICANT: Daniel, Muller K
; TITLE OF INVENTION: Human Bikunin
; FILE REFERENCE: 96-223-ZZ
; CURRENT APPLICATION NUMBER: US/09/974,026
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 09/144,428
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: PCT/US97/03894
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: US 08/725,251
; PRIOR FILING DATE: 1996-10-04
; PRIOR APPLICATION NUMBER: US 60/019,793
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: US 60/013,106
; PRIOR FILING DATE: 1996-03-11
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 45
```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide used in in vitro mutagenesis in Example 9.
US-09-974-026-61

Query Match          15.3%; Score 19; DB 12; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 ACGTGGCCATCTGCAGGG 23
        |||||||
DB       38 ACGTGGCCATCTGCAGGG 20

RESULT 13
US-10-027-632-113318/c
; Sequence 113318, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113318
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113318

Query Match          15.3%; Score 19; DB 12; Length 3372;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      70 AGAGGACTGTGTTGGTGC 88
        |||||||
DB       1339 AGAGGACTGTGTTGGTGC 1321

RESULT 14
US-10-027-632-113318/c
; Sequence 113318, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
```

;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 113318
;; LENGTH: 3372
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-113318

Query Match 15.3%; Score 19; DB 13; Length 3372;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 AGAGGACTGTGTGTCGTC 88
DB 1339 AGAGGACTGTGTGTCGTC 1321

RESULT 15
US-09-918-995-19059/C
;; Sequence 19059, Application US/09918995
;; Publication No. US20030073623A1
;; GENERAL INFORMATION:
;; APPLICANT: HySeq, Inc.
;; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;; FILE REFERENCE: 20411-756
;; CURRENT APPLICATION NUMBER: US/09/918,995
;; PRIOR FILING DATE: 2001-07-30
;; PRIOR APPLICATION NUMBER: US/09/235,076
;; PRIOR FILING DATE: 1999-01-20
;; NUMBER OF SEQ ID NOS: 38054
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 19059
;; LENGTH: 336
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc.feature
;; LOCATION: (1)...(336)
;; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-19059

Query Match 14.5%; Score 18; DB 11; Length 336;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 GGGAGATGTATCATTC 108
DB 135 GGGAGATGTATCATTC 118

RESULT 16
US-09-998-598-450/C
;; Sequence 450, Application US/0998598
;; Patent No. US20020150922A1
;; GENERAL INFORMATION:
;; APPLICANT: Stolk, John A.
;; APPLICANT: Xu, Jiangchun
;; APPLICANT: Chenuault, Ruth A.
;; APPLICANT: Meagher, Madelein Joy
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;; FILE REFERENCE: 210121.561
;; CURRENT APPLICATION NUMBER: US/09/998,598
;; PRIOR FILING DATE: 2001-11-16

;; NUMBER OF SEQ ID NOS: 2606
;; SOFTWARE: Corixa Invention Disclosure Database
;; SEQ ID NO 450
;; LENGTH: 541
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-998-598-450

Query Match 14.5%; Score 18; DB 10; Length 541;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 GGGAGATGTATCATTC 108
DB 154 GGGAGATGTATCATTC 137

RESULT 17
US-10-027-632-31061
;; Sequence 31061, Application US/10027632
;; Publication No. US20030204075A9
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; PRIOR FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 31061
;; LENGTH: 930
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-31061

Query Match 14.5%; Score 18; DB 12; Length 930;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 ATTCTTCACAACTGCT 122
DB 774 ATTCTTCACAACTGCT 791

RESULT 18
US-10-027-632-31061
;; Sequence 31061, Application US/10027632
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; PRIOR FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31061
LENGTH: 930
TYPE: DNA
ORGANISM: Human
US-10-027-632-31061

Query Match 14.5%; Score 18; DB 13; Length 930;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 ATTCCTCCCAACTGCT 122
|||||
DB 774 ATTCCTCCCAACTGCT 791

RESULT 19

US-10-106-698-799/c
Sequence 799, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005PI
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 799
LENGTH: 2630
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1676)..(1676)
OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-799

Query Match 14.5%; Score 18; DB 14; Length 2630;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 GGGAGATGTAAATTC 108
|||||
DB 1143 GGGAGATGTAAATTC 1126

RESULT 20

US-10-027-632-276727/c
Sequence 276727, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 276727
LENGTH: 252
TYPE: DNA
ORGANISM: Human
US-10-027-632-276727

Query Match 13.7%; Score 17; DB 12; Length 252;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GCAGGTCAGGTGATG 34
|||||
DB 51 GCAGGTCAGGTGATG 35

RESULT 21

US-10-027-632-276727/c
Sequence 276727, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 276727
LENGTH: 252
TYPE: DNA
ORGANISM: Human
US-10-027-632-276727

Query Match 13.7%; Score 17; DB 13; Length 252;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GCAGGTCAGGTGATG 34
|||||
DB 51 GCAGGTCAGGTGATG 35

RESULT 22
US-09-864-761-4974
Sequence 4974, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmics-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 4974
LENGTH: 355
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL031056.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
US-09-864-761-4974

Query Match 13.7%; Score 17; DB 9; Length 355;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22 GGTCCAGGTGATGATG 38
|||
Db 281 GGTCCAGGTGATGATG 297

RESULT 23
US-09-864-761-21703
Sequence 21703, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmics-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21703
LENGTH: 355
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL031056.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
OTHER INFORMATION: NT HIT: AB019786.1, EVALUE 1.20e-02
OTHER INFORMATION: EST_HUMAN HIT: AA495851.1, EVALUE 2.20e-01
US-09-864-761-21703

Query Match 13.7%; Score 17; DB 9; Length 355;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 GGTCCAGGTGATGATG 38
DB 281 GGTCCAGGTGATGATG 297

RESULT 24
US-10-066-543-1656
Sequence 1656, Application US/10066543
Publication No. US20030087818A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jianshun
APPLICANT: Indrias, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066,543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1656
LENGTH: 464
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 211, 453, 454
OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1656

Query Match 13.7%; Score 17; DB 14; Length 464;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CCATCTGCAGGGTCCAG 28
DB 433 CCATCTGCAGGGTCCAG 449

RESULT 25
US-09-918-995-31855/C
Sequence 31855, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31855

LENGTH: 500
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(500)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31855

Query Match 13.7%; Score 17; DB 11; Length 500;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CCATCTGCAGGGTCCAG 28
DB 482 CCATCTGCAGGGTCCAG 466

RESULT 26
US-10-027-632-145248/C
Sequence 145248, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 145248
LENGTH: 712
TYPE: DNA
ORGANISM: Human
US-10-027-632-145248

Query Match 13.7%; Score 17; DB 12; Length 712;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 ATCTGCAGGGTCCAGT 30
DB 689 ATCTGCAGGGTCCAGT 673

RESULT 27
US-10-027-632-145248/C
Sequence 145248, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676

```

; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145248
; LENGTH: 712
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-145248
```

```
Query Match      13.7%; Score 17; DB 13; Length 712;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      14 ATCTGAGGTCACAGT 30
      |||||
Db      689 ATCTGAGGTCACAGT 673
```

```

RESULT 28
US-10-027-632-162982
; Sequence 162982, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162982
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162982
```

```
Query Match      13.7%; Score 17; DB 12; Length 833;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      50 TGTCAAGCTGAACAA 66
      |||||
Db      194 TGTCAAGCTGAACAA 210
```

```

RESULT 29
US-10-027-632-162983
; Sequence 162983, Application US/10027632
```

```

; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162983
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162983
```

```
Query Match      13.7%; Score 17; DB 12; Length 833;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      50 TGTCAAGCTGAACAA 66
      |||||
Db      194 TGTCAAGCTGAACAA 210
```

```

RESULT 30
US-10-027-632-162982
; Sequence 162982, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162982
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162982
```

```
Query Match      13.7%; Score 17; DB 13; Length 833;
Best Local Similarity 100.0%; Pred. No. 19;
```


Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 TGTCAAGCTGAACA 66
|||||
Db 194 TGTCAAGCTGAACA 210

RESULT 31
US-10-027-632-162983
; Sequence 162983, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 162983
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162983

Query Match 13.7%; Score 17; DB 13; Length 833;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TGTCAAGCTGAACA 66
|||||
Db 194 TGTCAAGCTGAACA 210

RESULT 32
US-10-027-632-255720
; Sequence 255720, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 255720
; LENGTH: 1781
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-255720

Query Match 13.7%; Score 17; DB 12; Length 1781;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TTCCTTCACACTGCT 122
|||||
Db 1695 TTCCTTCACACTGCT 1711

RESULT 33
US-10-027-632-255720
; Sequence 255720, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 255720
; LENGTH: 1781
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-255720

Query Match 13.7%; Score 17; DB 13; Length 1781;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TTCCTTCACACTGCT 122
|||||
Db 1695 TTCCTTCACACTGCT 1711

RESULT 34
US-10-281-904-3/c
; Sequence 3, Application US/10281904
; Publication No. US20030119036A1
; GENERAL INFORMATION:
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: METHODS OF USING 48149, A HUMAN
; FILE REFERENCE: MP101-174P1M
; CURRENT APPLICATION NUMBER: US/10/281,904
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: 60/335,084
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 3
; LENGTH: 2904
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-281-904-3

Query Match
Best Local Similarity 13.7%; Score 17; DB 14; Length 2904;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 CCATCTGCAGGGTCCAG 28
Db 1642 CCATCTGCAGGGTCCAG 1626

RESULT 35
US-09-880-107-2295/c
; Sequence 2295, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2295
; LENGTH: 3494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M22324
US-09-880-107-2295

Query Match
Best Local Similarity 13.7%; Score 17; DB 10; Length 3494;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 CCATCTGCAGGGTCCAG 28
Db 1762 CCATCTGCAGGGTCCAG 1746

RESULT 36
US-10-205-823-31/c
; Sequence 31, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ganmavarrapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatc, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
```

```

; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 3494
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-31

Query Match
Best Local Similarity 13.7%; Score 17; DB 14; Length 3494;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 CCATCTGCAGGGTCCAG 28
Db 1762 CCATCTGCAGGGTCCAG 1746

RESULT 37
US-10-264-374-200/c
; Sequence 200, Application US/10264374
; Publication No. US20030113320A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; FILE REFERENCE: P-LJ 3203
; CURRENT APPLICATION NUMBER: US/10/264,374
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US/09/139,802
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 08/926,914
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 08/710,067
; PRIOR FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 200
; LENGTH: 3494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(3024)
US-10-264-374-200

Query Match
Best Local Similarity 13.7%; Score 17; DB 14; Length 3494;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 CCATCTGCAGGGTCCAG 28
Db 1762 CCATCTGCAGGGTCCAG 1746

RESULT 38
US-10-281-904-1/c
; Sequence 1, Application US/10281904
; Publication No. US20030119036A1
; GENERAL INFORMATION:
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: METHODS OF USING 48149, A HUMAN
```

```
/ TITLE OF INVENTION: AMINOPEPTIDASE FAMILY MEMBER
/ FILE REFERENCE: MP101-1741RM
/ CURRENT APPLICATION NUMBER: US/10/281,904
/ CURRENT FILING DATE: 2002-10-28
/ PRIOR APPLICATION NUMBER: 60/335,084
/ PRIOR FILING DATE: 2001-10-31
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 3494
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-281-904-1

Query Match      13.7%; Score 17; DB 14; Length 3494;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      12 CCATCTGCAGGGTCCAG 28
Db      1762 CCATCTGCAGGGTCCAG 1746

RESULT 39
US-10-133-013-221/c
/ Sequence 221, Application US/10133013
/ Publication No. US20030166903A1
/ GENERAL INFORMATION:
/ APPLICANT: Astrimoff, Anna
/ APPLICANT: Bandman, Olga
/ APPLICANT: Cocks, Benjamin G.
/ TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE
/ FILE REFERENCE: PA-0049 US
/ CURRENT APPLICATION NUMBER: US/10/133,013
/ CURRENT FILING DATE: 2002-04-25
/ PRIOR APPLICATION NUMBER: 60/287,067
/ PRIOR FILING DATE: 2001-04-27
/ NUMBER OF SEQ ID NOS: 271
/ SOFTWARE: PERL Program
/ SEQ ID NO 221
/ LENGTH: 3681
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20030166903A1 1136923.19
US-10-133-013-221

Query Match      13.7%; Score 17; DB 12; Length 3681;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      12 CCATCTGCAGGGTCCAG 28
Db      1960 CCATCTGCAGGGTCCAG 1944

RESULT 40
US-09-981-353-121/c
/ Sequence 121, Application US/09981353
/ Patent No. US20020160382A1
/ GENERAL INFORMATION:
/ APPLICANT: Lasek, Amy W.
/ APPLICANT: Jones, David A.
/ TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
/ FILE REFERENCE: PA-0038 US
/ CURRENT APPLICATION NUMBER: US/09/981,353
/ CURRENT FILING DATE: 2001-10-11
/ NUMBER OF SEQ ID NOS: 194
/ SOFTWARE: PERL Program
/ SEQ ID NO 121
/ LENGTH: 3686
/ TYPE: DNA

/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20020160382A1 1344279CB1
US-09-981-353-121

Query Match      13.7%; Score 17; DB 10; Length 3686;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy      12 CCATCTGCAGGGTCCAG 28
Db      1962 CCATCTGCAGGGTCCAG 1946

RESULT 41
US-10-044-090-218/c
/ Sequence 218, Application US/10044090
/ Publication No. US20020137081A1
/ GENERAL INFORMATION:
/ APPLICANT: Olga Bandman
/ TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
/ FILE REFERENCE: PA-0028 US
/ CURRENT APPLICATION NUMBER: US/10/044,090
/ CURRENT FILING DATE: 2002-01-09
/ NUMBER OF SEQ ID NOS: 850
/ SOFTWARE: PERL Program
/ SEQ ID NO 218
/ LENGTH: 3686
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20020137081A1 1344279CB1
US-10-044-090-218

Query Match      13.7%; Score 17; DB 13; Length 3686;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      12 CCATCTGCAGGGTCCAG 28
Db      1962 CCATCTGCAGGGTCCAG 1946

RESULT 42
US-10-084-817-86/c
/ Sequence 86, Application US/10084817
/ Publication No. US20030119009A1
/ GENERAL INFORMATION:
/ APPLICANT: Susan Stuart
/ APPLICANT: Jed G. Nuchtern
/ APPLICANT: Sharon E. Pion
/ APPLICANT: Jason M. Shohet
/ TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
/ FILE REFERENCE: PA-0046 US
/ CURRENT APPLICATION NUMBER: US/10/084,817
/ CURRENT FILING DATE: 2002-02-25
/ PRIOR APPLICATION NUMBER: 60/270,784
/ PRIOR FILING DATE: 2001-02-23
/ NUMBER OF SEQ ID NOS: 365
/ SOFTWARE: PERL Program
/ SEQ ID NO 86
/ LENGTH: 3686
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20030119009A1 1344279CB1
US-10-084-817-86

Query Match      13.7%; Score 17; DB 14; Length 3686;
Best Local Similarity 100.0%; Pred. No. 19;
```

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CCATCTGCAGGCTCCAG 28
Db 1962 CCATCTGCAGGCTCCAG 1946

RESULT 43

US-09-925-297-338/c
; Sequence 338, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 338
; LENGTH: 3728
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3707)
; OTHER INFORMATION: n equals a,t,g, or c
; LOCATION: (3713)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-338

Query Match 13.7%; Score 17; DB 9; Length 3728;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CCATCTGCAGGCTCCAG 28
Db 1972 CCATCTGCAGGCTCCAG 1956

RESULT 44

US-10-133-013-222/c
; Sequence 222, Application US/10133013
; Publication No. US20030166903A1
; GENERAL INFORMATION:
; APPLICANT: Astromoff, Anna
; APPLICANT: Bandman, Olga
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE
; FILE REFERENCE: PA-0049 US
; CURRENT APPLICATION NUMBER: US/10/133,013
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/287,067
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PRL Program
; SEQ ID NO 222
; LENGTH: 3815
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030166903A1 1136923.18
US-10-133-013-222

Query Match 13.7%; Score 17; DB 12; Length 3815;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CCATCTGCAGGCTCCAG 28
Db 1961 CCATCTGCAGGCTCCAG 1945

RESULT 45

US-09-764-891-6245/c
; Sequence 6245, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6245
; LENGTH: 11150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-6245

Query Match 13.7%; Score 17; DB 11; Length 11150;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CCATCTGCAGGCTCCAG 28
Db 76 CCATCTGCAGGCTCCAG 60

RESULT 46

US-10-205-428-561/c
; Sequence 561, Application US/10205428
; Publication No. US20030108907A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA117C1
; CURRENT APPLICATION NUMBER: US/10/205,428
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1019
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 561
; LENGTH: 11150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-428-561

Query Match 13.7% Score 17; DB 14; Length 11150;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CCATCTGCAGGCTCCAG 28
DB 76 CCATCTGCAGGCTCCAG 60

RESULT 47
US-09-878-574-8350
Sequence 8350, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 8350
LENGTH: 253
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: 701101059H1
US-09-878-574-8350

Query Match 12.9% Score 16; DB 10; Length 253;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TGAACAACAACAAGAG 73
DB 115 TGAACAACAACAAGAG 130

RESULT 48
US-09-682-706-1/C
Sequence 1, Application US/09682706
Patent No. US20020082409A1
GENERAL INFORMATION:
APPLICANT: Hsu, Sheau-Yu
APPLICANT: Hsieh, Aaron
TITLE OF INVENTION: Stresscopins and their ses
FILE REFERENCE: STAN210
CURRENT APPLICATION NUMBER: US/09/682,706
CURRENT FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/276,615
PRIOR FILING DATE: 2001-03-15
PRIOR APPLICATION NUMBER: 60/244,128
PRIOR FILING DATE: 2000-10-26
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 339
TYPE: DNA
ORGANISM: Homo Sapiens
US-09-682-706-1

Query Match 12.9% Score 16; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 GTGCTGTGGGAGAAAT 98
DB 116 GTGCTGTGGGAGAAAT 101

RESULT 49
US-09-983-965-5910/C
Sequence 5910, Application US/09983965
Patent No. US20020137160A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengping
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 37-21(10297)C
CURRENT APPLICATION NUMBER: US/09/983,965
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: US 09/465,231
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: US 60/113,678
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 5912
SEQ ID NO 5910
LENGTH: 397
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
OTHER INFORMATION: Clone ID: 64-LIB34-071-Q1-E1-H8
US-09-983-965-5910

Query Match 12.9% Score 16; DB 10; Length 397;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGGCTCCAGGTGATGG 35
DB 88 AGGCTCCAGGTGATGG 73

RESULT 50
US-09-919-473-1/C
Sequence 1, Application US/09919473
Patent No. US20020127221A1
GENERAL INFORMATION:
APPLICANT: Vale, Wylie Walker Jr.
APPLICANT: Lewis, Kathy Ann
APPLICANT: Reyes, Teresa Marie
APPLICANT: Hogenesch, John Beren
APPLICANT: Sawchenko, Paul Emil
APPLICANT: Vaughan, Joan Maureen
APPLICANT: Rivier, Jean Edouard Frederic
TITLE OF INVENTION: Urocortin Proteins and Uses Thereof
FILE REFERENCE: D6334
CURRENT APPLICATION NUMBER: US/09/919,473
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/273,969
PRIOR FILING DATE: 2001-03-07
NUMBER OF SEQ ID NOS: 13
SEQ ID NO 1
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 7..345
OTHER INFORMATION: DNA Sequence encoding human
OTHER INFORMATION: Urocortin-related peptide (hURP)
US-09-919-473-1

Query Match 12.9% Score 16; DB 10; Length 399;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 GTGCTGTGGGAGAAAT 98
DB 116 GTGCTGTGGGAGAAAT 101

Fri Nov 7 15:17:33 2003

us-09-509-779-3_copy_141_264.rnpb

Page 16

Db 122 GTGCTCTGGGAGAAAT 107

Search completed: November 7, 2003, 13:55:40
Job time : 479.215 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 02:39:56 ; Search time 126.47 Seconds
(without alignments)
2646.715 Million cell updates/sec

Title: US-09-509-779-3_COPY_141_264

Perfect score: 124

Sequence: 1 CGATACGTGCGCATCTGCA.....ATTCTTCACAACTGCTGC 124

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_19jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	342	21	AAA96883
2	124	100.0	342	22	AAH43570
3	124	100.0	342	24	AAH39674
4	124	100.0	754	20	AAH87325
5	124	100.0	754	20	AAH87330
6	124	100.0	754	20	AAH87331
7	124	100.0	754	20	AAH87314
8	124	100.0	836	22	AAH25860

9	124	100.0	836	22	AAH26319	Human cDNA encodin
10	124	100.0	836	25	ABX73201	Human novel polynu
11	124	100.0	836	25	ABX73660	Human novel polynu
12	124	100.0	962	21	AAH77504	Human OREF ORF3059
13	121	97.6	754	20	AAH87324	Human sensitive to
14	116	93.5	754	20	AAH87317	Human sensitive to
15	112	90.3	754	20	AAH87323	Human sensitive to
16	107	86.3	754	20	AAH87318	Human sensitive to
17	107	86.3	754	20	AAH87319	Human sensitive to
18	103	83.1	754	20	AAH87322	Human sensitive to
19	103	83.1	754	20	AAH87326	Human sensitive to
20	97	78.2	754	20	AAH87321	Human sensitive to
21	90	72.6	747	20	AAH87315	Human sensitive to
22	89	71.8	1152	22	AAH25847	Human apoptosis as
23	83	66.9	754	20	AAH87320	Human sensitive to
24	76	61.3	754	20	AAH87328	Human sensitive to
25	74	59.7	754	20	AAH87327	Human sensitive to
26	74	59.7	754	20	AAH87329	Human sensitive to
27	43	34.7	224	20	AAH41085	Human secreted pro
28	42	33.9	706	20	AAH87316	Human secreted pro
29	41	33.1	441	21	AAH77493	Human OREF ORF3048
30	35	28.2	60	24	ABH40537	Human spliced tran
31	35	28.2	439	24	ABH21414	Human polynucleoti
32	32	25.8	1140	20	AAH87313	Mouse sensitive to
33	32	19.4	264	22	AAH16209	Human breast cance
34	24	19.4	596	22	AAH25052	Human breast cance
35	24	19.4	630	22	AAH22239	Human breast cance
36	19	15.3	45	21	AAH70390	Site directed muta
37	18	14.5	18	20	AAH87337	Human sensitive to
38	18	14.5	349	21	AAH29980	Human secreted pro
39	18	14.5	541	24	ABH87139	Human colon cancer
40	18	14.5	2405	15	AAH92962	Protein kinase (CK
41	18	14.5	2405	16	AAH92962	Human HR235-like c
42	18	14.5	2630	22	AAH33733	Human colon cancer
43	17	13.7	239	23	AAH84630	DNA encoding novel
44	17	13.7	355	22	ABA46329	Human breast cell
45	17	13.7	355	22	ABA51431	Human breast cell

ALIGNMENTS

RESULT 1

ID AAA96883 standard; DNA, 342 BP.

AC AAA96883;

DT 19-FEB-2001 (first entry)

DE Nucleotide sequence of human ring finger protein ROC2.

KW ROC2; cullin; ring finger protein; APC1; APC complex; SCF pathway;

KW cullin dependent ubiquitin ligase; CDK inhibitor Sic1 degradation;

KW tumor; ss.

OS Homo sapiens.

XX Key

XX CDS

XX FT

XX FT

XX FT

XX FT

XX FT

Location/Qualifiers
1..342
/tag= a
/product= "ROC2"
/transl_except= (pos: 28..30, aa: Pro)
/transl_except= (pos: 34..36, aa: Val)
/transl_except= (pos: 40..42, aa: Ser)
/transl_except= (pos: 58..60, aa: Ala)
/transl_except= (pos: 67..69, aa: Lys)
/transl_except= (pos: 109..111, aa: Ala)

WO200058472-A2.

05-OCT-2000.

XX

PF 31-MAR-2000; 2000OWO-US08592.
 XX
 PR 31-MAR-1999; 99US-0127261.
 PR 22-NOV-1999; 99US-0166927.
 XX
 PA (UNYCN-) UNIV NORTH CAROLINA.
 XX
 XX
 PI Xiong Y, Ohta T;
 XX
 DR WPI; 2000-647235/62.
 DR P-PSDB; AAB19161.
 XX
 PT Novel nucleic acid encoding cullin regulating ring finger proteins,
 PT termed as ROC proteins similar to anaphase-promoting complex 11, for
 PT therapeutic and diagnostic use -
 XX
 XX
 PS Claim 18; Fig 2B; 83pp; English.
 XX
 CC The present sequence encodes a human ROC2 ring finger protein. The
 CC specification also describes human ROC1, ROC1 and ROC2 are similar
 CC to APC11, a subunit of the APC complex. The proteins stimulate cullin
 CC dependent ubiquitin ligase activity. ROC1 functions in vivo as an
 CC essential regulator of CDK inhibitor Sic1 degradation by the SCF
 CC (undefined) pathway. ROC proteins are useful for screening bioactive
 CC agents that interfere with the binding of ROC proteins with cullin
 CC proteins. Pharmaceutical formulations comprising ROC proteins are
 CC useful for diagnostic and therapeutic purposes, preferably for
 CC diagnosing and treating tumours.
 XX
 XX Sequence 342 BP; 82 A; 84 C; 108 G; 68 T; 0 other;

XX Issakani SD, Huang J, Sheung J, Pray TR;
PI WPI: 2001-626445/72.
XX P-PSDB; AAB47600.
DR
XX
XX
PT Assaying ubiquitin ligase activity for identifying modulators of
PT ubiquitination, by combining ubiquitin, ubiquitin activating
PT conjugating enzyme, ubiquitin ligase and measuring amount of ubiquitin
PT bound to the ligase -
XX
XX
XX Example 1; Fig 12A; 98pp; English.

The sequences given in AAB43568-72 encode proteins which may be used in the method of the invention for assaying ubiquitin ligase activity. The method comprises combining under conditions that favour ubiquitin ligase activity, tag1-ubiquitin, E1 (ubiquitin activating enzyme), E2 (ubiquitin conjugating enzyme) and E3 (ubiquitin ligase) and measuring the amount of tag1-ubiquitin bound to the E3. The method is useful for assaying ubiquitin ligase activity and ubiquitination enzyme activity which is useful for identifying ubiquitination modulator. The method comprises combining tag1-ubiquitin, the modulator, E1, E2 and tag2-E3 and measuring the amount of tag1-ubiquitin bound to tag2-E3 or combining tag1-ubiquitin, modulator, E1 and tag3-E2 and measuring the amount of tag1-ubiquitin bound to tag3-E2. Ubiquitin ligase activity is measured directly, eliminating the need for target proteins and subsequent analysis such as separating ligated from unligated material in an SDS-PAGE procedure. This allows multi-well array analysis and high throughput screening techniques for modulators of ubiquitination activity. The method also allows the analysis of many different

Query Match	100.0%	Score 124	DB 21	Length 342
Best Local Similarity	100.0%	Pred. No. 1e-55		
Matches 124	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY 1 CGATTCGCGCCATCTGACGGGTCCAGGTGATGATGCTCTTAAATGTCAAGCTGA 60

CC measurement of ligated ubiquitin.
 CC directly or indirectly and this allows for easy and rapid detection and
 CC prior identification of specific target substrates. Ubiquitin is labeled
 CC combinations of *ε*-E3 components and *ε*-E3 combinations without requiring
 CC combinations of *ε*-E3 components and *ε*-E3 combinations without requiring
 XX
 XX
 SQ Sequence 342 BP; 82 A; 84 C; 108 G; 68 T; 0 other;

QY 1 CGATACCTGGCCATCTGCAGAGGTCAGGATGATGATGCTCTTAAATGTCAGAGTGA 60
Db CGATACCTGGCCATCTGCAGAGGTCAGGATGATGATGCTCTTAAATGTCAGAGTGA 200
QY 61 AAACAACAGAGGACTGTGTGTGCTCGGGAGAAATGATATATCTCTCCCAACTG 120
Db AAACAACAGAGGACTGTGTGTGCTCGGGAGAAATGATATATCTCTCCCAACTG 260
QY 121 CTGC 124
Db 261 CTGC 264

SQ	Sequence	342 BP;	82 A;	84 C;	108 G;	68 T;	0 other;
	Query Match		100.0%;	Score 124;	DB 22;	Length 342;	
	Best Local Similarity		100.0%;	Pred. No. 1e-55;			
	Matches 124;	Conservative	0;	Mismatches	0;	Indels	0; Gaps 0;
QY	1	CGATACCTGGGCCCATCTGCAGAGGTCACAGTGTATGGATCCGTCTTTAGATGTCAAGCTGA	60				
Dd	141	CGATACCTGGGCCCATCTGCAGAGGTCACAGTGTATGGATCCGTCTTTAGATGTCAAGCTGA	200				
OY	61	AACAAACAAGAGAGCATTGTCTTGCTGGGAGGAAATTAATATATCTTCCACAAATCG	120				

RESULT 2	
AAH43570	
ID	AAH43570 standard; cDNA, 342 bp
XX	
AC	AAH43570;
XX	
DT	07-JUN-2002 (first entry)
XX	
DE	ROC2 coding sequence.

QY	121	CTGC	124
Db	261	CTGC	264

RESULT 3
AAD39674
AAD39674 standard; DNA: 342 BP.

km Assay; ubiquitin ligase; teg1-ubiquitin; E1; E2;
 km ubiquitin activating enzyme; ubiquitin conjugating enzyme; E3;
 km ubiquitin ligase; ubiquitination modulator; ss.
 xx
 Unidentified.
 OS

AC	AAD39674,
XX	
DT	22-OCT-2002 (first entry)
XX	
HE	Human RING finger protein, ROC2 DNA.

PN	WO200175145-A2.
XX	
PD	11-OCT-2001.

Ubiquitin ligase; U1, tag1-ubiquitin; E1, ubiquitin-activating enzyme; E2, ubiquitin-conjugating enzyme/ubiquitin carrier protein; E3, human; RING finger protein; gene: ds.

PF 03-APR-2001; 2001WO-US10906.
XX
PR 03-APR-2000; 2000US-0542497.
VV

OS	Homo sapiens.
XX	
TH	Key
TH	Location/Qualifiers

PA (RIGE-) RIGEL PHARM INC.

ET	1.342
ET	/*tga=
CDS	a

FT XX /product= "Human ROC2 protein"
PN XX US2002042083-A1.
XX
PD 11-APR-2002.
XX
XX 03-APR-2001; 2001US-0826312.
XX PF
XX 03-APR-2000; 2000US-0542497.
XX PR
XX (RIGF-) RIGEL PHARM INC.
XX
XX Issakani SD, Huang J, Sheung J, Pray TR;
PI WPI; 2002-488718/52.
XX
XX P-PSDB; AAE24621.
DR
PT Assay for ubiquitin ligase activity, useful for identifying modulators,
PT by measuring binding of labeled ubiquitin to ubiquitin ligase
XX
XX
XX PS Disclosure; Fig 12A; 56pp; English.
XX
XX CC The invention relates to an assay for ubiquitin ligase (UL) activity
CC which comprises incubating tag1-ubiquitin, E1 (ubiquitin-activating
CC enzyme), E2 (ubiquitin-conjugating enzyme/ubiquitin carrier protein)
CC and E3 (UL) and measuring the amount of tag1-ubiquitin bound to E3. The
CC method is particularly used to screen for modulators of UL activity. The
CC present sequence is human RING finger protein, ROC2 DNA.
XX
XX SQ Sequence 342 BP; 82 A; 84 C; 108 G; 68 T; 0 other;
XX
XX Query Match 100.0%; Score 124; DB 24; Length 342;
XX Best Local Similarity 100.0%; Pred. No. 1e-55;
XX Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CGATACGTGCGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAACTGA 60
XX 141 CGATACGTGCGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAACTGA 200
XX DB
XX QY 61 AAACAACAAGAGACTGTGTGTGCTGTGGGAGATGATTAATCATTCCTTCCACAAC 120
XX 201 AAACAACAAGAGACTGTGTGTGCTGTGGGAGATGATTAATCATTCCTTCCACAAC 260
XX DB
XX QY 121 CTGC 124
XX 261 CTGC 264
XX DB
XX
XX RESULT 4
XX AA87325
XX ID AA87325 standard; cDNA; 754 BP.
XX
XX AC AA87325;
XX
XX DT 27-SEP-1999 (first entry)
XX
XX DE Human sensitive to apoptosis (SAG) gene mutant MM9.
XX
XX KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnerable; therapy; mutant; ds.
XX
XX OS Homo sapiens.
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT 1..342
XX FT /*tag= a
XX FT replace(295,T)
XX FT /*tag= b
XX FT /note= "C99S mutation"
XX
XX PN WO932514-A2.

XX
XX PD 01-JUL-1999.
XX
XX PF 15-DEC-1998; 98WO-US26705.
XX
XX PR 11-SEP-1998; 98US-0099840.
XX PR 19-DEC-1997; 97US-0068179.
XX
XX PA (WARN) WARNER LAMBERT CO.
XX
XX PI Sun Y;
XX
XX DR WPI; 1999-430152/36.
XX P-PSDB; AA06503.
XX
XX PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
XX promoting cell growth and protecting cells against apoptosis
XX
XX PS Claim 15; Page 72-73; 84pp; English.
XX
XX CC This is the nucleotide sequence of human sensitive to apoptosis
XX (SAG) mutant gene MM9, which codes for a SAG protein (see AA06503)
XX in which the Cys residue at position 99 of the native protein (see
XX CC AA06492) is replaced by a Ser residue owing to a mutation of codon
XX 99 from TGC to AGC obtained by site-directed mutagenesis of SAG
XX cDNA. This residue is in zinc-ring finger 2 of SAG. Single and
XX double SAG mutants (see AA87317-31) were made in order to determine
XX the role of each cysteine residue of SAG in haem binding and SAG
XX oligomerization. These properties were unaffected by the MM9
XX mutation. SAG is a novel zinc finger protein that promotes
XX cell growth, protects cells from apoptosis, scavenges oxygen
XX radicals and can be used for the reversion of a tumour phenotype.
XX SAG genes, and mutant SAG genes, can be used to protect cells from
XX apoptosis induced by redox reagents. They can also be used for the
XX recombinant production of SAG proteins, which are molecular targets
XX in the development of drugs against neurodegenerative disorders,
XX cancers and muscle dystrophy, and promoting wound healing.
XX
XX SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
XX
XX Query Match 100.0%; Score 124; DB 20; Length 754;
XX Best Local Similarity 100.0%; Pred. No. 1e-55;
XX Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CGATACGTGCGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAACTGA 60
XX 141 CGATACGTGCGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAACTGA 200
XX DB
XX QY 61 AAACAACAAGAGACTGTGTGTGCTGTGGGAGATGATTAATCATTCCTTCCACAAC 120
XX 201 AAACAACAAGAGACTGTGTGTGCTGTGGGAGATGATTAATCATTCCTTCCACAAC 260
XX DB
XX QY 121 CTGC 124
XX 261 CTGC 264
XX DB
XX
XX RESULT 5
XX AA87330
XX ID AA87330 standard; cDNA; 754 BP.
XX
XX AC AA87330;
XX
XX DT 27-SEP-1999 (first entry)
XX
XX DE Human sensitive to apoptosis (SAG) gene mutant MM14.
XX
XX KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnerable; therapy; mutant; ds.
XX
XX OS Homo sapiens.
XX OS Synthetic.

XX	Key	Location/Qualifiers
XX	FT CDS	1..342
XX	FT mutation	/*tag= a replace(295,C)
XX	FT mutation	/*tag= b /note= "C99S mutation"
XX	FT mutation	replace(304,C) /*tag= c /note= "C102S mutation"
XX	PN	WO9932514-A2.
XX	PD	01-JUL-1999.
XX	XX	
XX	PF	15-DEC-1998; 98WO-US26705.
XX	PR	11-SEP-1998; 98US-0099840.
XX	PR	19-DEC-1997; 97US-0068179.
XX	PA	(WARN) WARNER LAMBERT CO.
XX	PI	Sun Y;
XX	DR	WPI; 1999-430152/36.
XX	DR	P-PSDB; AA06508.
XX	PT	SAG: Sensitive to Apoptosis Gene and related proteins, useful for promoting cell growth and protecting cells against apoptosis
XX	PS	Claim 15; Page 81; 84pp; English.
XX	CC	This is the nucleotide sequence of human sensitive to apoptosis (SAG) mutant gene MM14, which codes for a SAG protein (see AA06508) in which the Cys residues at positions 99 and 102 of the native protein (see AA06492) are replaced by Ser residues owing to mutations of codons 99 and 102 through site-directed mutagenesis of SAG cDNA. These residues are in zinc ring finger 2 of SAG. Single double SAG mutants (see AA87317-31) were made in order to determine the role of each cysteine residue of SAG in haem binding and SAG oligomerization. Haem binding was unaffected in MM14, but no oligomerization was observed, indicating that these residues are important for intermolecular disulfide bond formation. SAG is a novel zinc finger protein that promotes cell growth, protects cells from apoptosis, scavenges oxygen radicals and can be used for the reversal of a tumour phenotype. SAG genes, and mutant SAG genes, can be used to protect cells from apoptosis induced by redox reagents. They can also be used for the recombinant production of SAG proteins, which are molecular targets in the development of drugs against neurodegenerative disorders, cancers and muscle dystrophy, and promoting wound healing.
XX	XX	Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;
QY	Query Match	100.0%; Score 124; DB 20; Length 754;
QY	Best Local Similarity	100.0%; Pred. No. 16-55;
QY	Matches 124; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	1	CGATAGCTGGCCATCTGCAGGGGTCCAGGAGATGATGCTGTCTTAGATGTCAAGCTGA 60
Db	141	CGATTAGCTGGCCATCTGCAGGGGTCCAGGAGATGATGCTGTCTTAGATGTCAAGCTGA 2000
QY	61	AAACAACAAGAGAGAGACTGTGTGTGTGTGTGGGAGAGATGTATCATCTTCCACAACCTG 1200
Db	201	AAACAACAAGAGAGAGACTGTGTGTGTGTGTGGGAGAGATGTATCATCTTCCACAACCTG 2600
QY	121	CTGC 124
QY	261	CTGC 264

ID	AAx87331 standard; cDNA: 754 BP.
XX	
AC	AAx87331;
XX	
DT	27-SEP-1999 (first entry)
XX	
DE	Human sensitive to apoptosis (SAG) gene mutant MM15.
XX	
KW	SAG gene; sensitive to apoptosis; human; cancer; tumour; neurodegenerative disease; muscular dystrophy; wound healing; vulnerary; therapy; mutant; ds.
XX	
OS	Homo sapiens.
XX	
FT	Synthetic.
XX	
Key	Location/Qualifiers
FT	CDS 1..342
FT	/+tag= a
FT	replac(139,C)
FT	/+tag= b
FT	/note= "C47S mutation"
XX	
PN	WO9352514-A2.
XX	
PD	01-JUL-1999.
XX	
PF	15-DEC-1998; 98WO-US26705.
XX	
PR	11-SEP-1998; 98US-0099840.
XX	
PR	19-DEC-1997; 97US-0068179.
XX	
PA	(WARN) WARNER LAMBERT CO.
XX	
PI	Sun Y;
XX	
DR	WPI: 1999-430152/36.
XX	
P-PSDB:	AAV06509.
XX	
PT	SAG: Sensitive to Apoptosis Gene and related proteins, useful for promoting cell growth and protecting cells against apoptosis
XX	
PS	Claim 15; Page 82-83; 84pp; English.
XX	
CC	This is the nucleotide sequence of human sensitive to apoptosis (SAG) mutant gene MM15, which codes for a SAG protein (see AAV06509) in which the Cys residue at position 47 of the native protein (see AAV06492) is replaced by a Ser residue owing to a mutation of codon 47 from TGC to AGC obtained by site-directed mutagenesis of SAG cDNA. This residue is in a GADPH binding site of SAG. Single and double SAG mutants (see AAx87317-31) were made to determine the role of each cysteine residue of SAG in haem binding and SAG oligomerization. These properties were unaffected by the MM15 mutation. SAG is a novel zinc finger protein that promotes cell growth, protects cells from apoptosis, scavenges oxygen radicals and can be used for the reversal of a tumour phenotype. SAG genes, and mutant SAG genes, can be used to protect cells from apoptosis induced by redox reagents. They can also be used for the recombinant production of SAG proteins, which are molecular targets in the development of drugs against neurodegenerative disorders, cancers and muscle dystrophy, and promoting wound healing.
XX	
XX	
SO	Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
Query Match	100.0%; Score 124; DB 20; Length 754;
Best Local Similarity	100.0%; Pred.No. 1e-55; Indels 0; Gaps 0
Matches 124; Conservative	0; Mismatches 0; Indels 0; Gaps 0
Dn	1 CGATACGTCGCCCATCTGCAGGGTCCAGGTGATGATGATGCTTTAAGTGCAAGCTGA 60 141 CGATACGTCGCCCATCTGCAGGGTCCAGGTGATGATGATGCTTTAAGTGCAAGCTGA 200 61 AAAACAACAAGAGACTGTGTTGTGTCTCGGGGAAGAATGATATCTTCTTCCAACAGC 120

PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
DR P-PSDB; MAU16332.
XX
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
XX
PS Claim 1; SEQ ID No 498; 980bp; English.
PS
XX
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis.
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi.
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention.

Query Match	100.0%	Score 124;	DB 22;	Length 836;
Best Local Similarity	100.0%	Pred. No. 1e-55;		
Matches 124;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGATACGCGCCCACTCTGACGGGATCCAGGTATGATGATCCGTCTTATGATGTCAGACCTGA	60	
Db	209	CGATACGCGCCCACTCTGACGGGATCCAGGTATGATGATCCGTCTTATGATGTCAGACCTGA	268	
Qy	61	AAACCAACAAAGAGCACTGTGTTGTGTCTGCGGAGATGTAATCATCTTCACCACTG	120	
Db	269	AAACCAACAAAGAGCACTGTGTTGTGTCTGCGGAGATGTAATCATCTTCACCACTG	328	
Qy	121	CTGC 124		
Db	329	CTGC 332		
RESULT 10				
ABX73201				
ID	ABX73201	standard; DNA; 836 BP.		
XX	ABX73201;			
AC	ABX73201;			
XX				
DT	18-MAR-2003	(first entry)		
XX				
DE	Human novel polynucleotide #29.			
XX				
KW	Human; gene; ds; neural disorder; immune system disorder; renal disorder;			
KW	muscular disorder; respiratory disease; reproductive disorder;			
KW	gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;			
KW	hyperproliferative disorder; inflammatory disease; allergic reaction;			
KW	blood related disorder; cancer; immunosuppressive; antiinflammatory;			
KW	cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;			
KW	haemostatic; antiarteriosclerotic.			
XX				
OS	Homo sapiens.			
XX				
PN	US2002132753-A1.			
PD				
XX	19-SEP-2002.			
PF				
XX	17-JAN-2001; 2001US-0764864.			
PR	31-JAN-2000; 2000US-179065P.			
PR	04-FEB-2000; 2000US-180628P.			
PR	28-JUN-2000; 2000US-214886P.			
PR	07-JUL-2000; 2000US-216647P.			
PR	07-JUL-2000; 2000US-216880P.			
PR	11-JUL-2000; 2000US-217487P.			
PR	11-JUL-2000; 2000US-217496P.			
PR	14-JUL-2000; 2000US-218290P.			
PR	26-JUL-2000; 2000US-220963P.			
PR	26-JUL-2000; 2000US-220964P.			
PR	14-AUG-2000; 2000US-224516P.			
PR	14-AUG-2000; 2000US-224519P.			
PR	14-AUG-2000; 2000US-225267P.			
PR	14-AUG-2000; 2000US-225268P.			
PR	14-AUG-2000; 2000US-225470P.			
PR	14-AUG-2000; 2000US-225474P.			
PR	14-AUG-2000; 2000US-225757P.			
PR	14-AUG-2000; 2000US-225758P.			
PR	22-AUG-2000; 2000US-225868P.			
PR	30-AUG-2000; 2000US-228924P.			
PR	01-SEP-2000; 2000US-229287P.			
PR	01-SEP-2000; 2000US-229343P.			
PR	01-SEP-2000; 2000US-229344P.			
PR	01-SEP-2000; 2000US-229345P.			
PR	05-SEP-2000; 2000US-229509P.			
PR	05-SEP-2000; 2000US-229513P.			
PR	08-SEP-2000; 2000US-231413P.			
PR	21-SEP-2000; 2000US-234223P.			
PR	21-SEP-2000; 2000US-234274P.			
PR	25-SEP-2000; 2000US-234977P.			

PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251869P.
PR 08-DEC-2000; 2000US-251869P.

XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.

PI Rosen CA, Ruben SM, Barash SC;

DR WPI; 2003-147444/14.

XX P-PSDB; ABUS4941.

PT New polypeptides and nucleic acids, useful in gene therapy for
PT treating, inhibiting or preventing e.g. neural, immune system,
PT muscular, respiratory, reproductive, gastrointestinal, pulmonary,
PT cardiovascular or renal disorders -

XX Claim 1; SEQ ID NO 39; 402bp; English.

XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid
CC arthritis and multiple sclerosis), muscular disorders, respiratory
CC diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis),
CC reproductive disorders, gastrointestinal disorders, pulmonary disorders,
CC cardiovascular disorders (e.g. congenital heart defects, Eisen's
CC anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute
CC kidney failure and end-stage renal disease), hyperproliferative disorders
CC (e.g. Hodgkin's disease and leukaemia), inflammatory diseases (e.g.
CC septic shock, bursitis and appendicitis), allergic reactions and
CC conditions (e.g. asthma), blood related disorders (e.g. thrombosis,
CC atherosclerosis and myocardial infarction) and cancerous diseases.
CC Sequences ABX73173-ABX74167 represent human novel polynucleotides of the
CC invention.
XX

SQ Sequence 836 BP; 264 A; 168 C; 209 G; 195 T; 0 other;

Query Match 100.0%; Score 124; DB 25; Length 836;

Best Local Similarity 100.0%; Pred. No. 1e-55; Mismatches 0; Gaps 0;

Matches 124; Conservative 0; Indels 0;

QY 1 CGATACGTGGCCATCTGACAGGTCCAGGTGATGATGCTGCTTAGATGTAACACTGA 60
DB 156 CGATACGTGGCCATCTGACAGGTCCAGGTGATGATGCTGCTTAGATGTAACACTGA 215
QY 61 AAAACAACAAGAGACTGTGTGTGTGTGTGTGGGAGATGTAATCATTCCTTCCACAACCTG 120
DB 216 AAAACAACAAGAGACTGTGTGTGTGTGTGTGGGAGATGTAATCATTCCTTCCACAACCTG 275

QY 121 CTGC 124
DB 276 CTGC 279

RESULT 11

ABX73660
ID ABX73660 standard; DNA; 836 BP.

XX ABX73660;

AC 18-MAR-2003 (first entry)

DE Human novel polynucleotide #488.

XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotoxic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.

XX Homo sapiens.

OS US2002132753-A1.

PD 19-SEP-2002.

PF 17-JAN-2001; 2001US-0764864.

XX 31-JAN-2000; 2000US-179065P.

XX 04-FEB-2000; 2000US-180628P.

XX 28-JUN-2000; 2000US-214886P.

XX 07-JUL-2000; 2000US-216647P.

XX 11-JUL-2000; 2000US-216880P.

XX 11-JUL-2000; 2000US-217487P.

XX 11-JUL-2000; 2000US-217496P.

XX 14-JUL-2000; 2000US-218290P.

XX 26-JUL-2000; 2000US-220963P.

XX 14-AUG-2000; 2000US-224518P.

XX 14-AUG-2000; 2000US-224519P.

XX 14-AUG-2000; 2000US-225267P.

XX 14-AUG-2000; 2000US-225268P.

XX 14-AUG-2000; 2000US-225270P.

XX 14-AUG-2000; 2000US-225447P.

XX 14-AUG-2000; 2000US-225757P.

XX 14-AUG-2000; 2000US-225758P.

XX 22-AUG-2000; 2000US-226868P.

XX 30-AUG-2000; 2000US-228924P.

XX 01-SEP-2000; 2000US-229287P.

XX 01-SEP-2000; 2000US-229343P.

XX 01-SEP-2000; 2000US-229344P.

XX 05-SEP-2000; 2000US-229345P.

XX 05-SEP-2000; 2000US-229509P.

XX 08-SEP-2000; 2000US-229513P.

XX 21-SEP-2000; 2000US-231413P.

XX 21-SEP-2000; 2000US-234223P.

XX 25-SEP-2000; 2000US-234374P.

XX 27-SEP-2000; 2000US-234977P.

XX 27-SEP-2000; 2000US-235834P.

XX 29-SEP-2000; 2000US-236327P.

XX 29-SEP-2000; 2000US-236367P.

XX 29-SEP-2000; 2000US-236368P.

XX 29-SEP-2000; 2000US-236369P.

XX 29-SEP-2000; 2000US-236370P.

XX 02-OCT-2000; 2000US-236802P.

XX 02-OCT-2000; 2000US-237037P.

XX 02-OCT-2000; 2000US-237038P.

XX 02-OCT-2000; 2000US-237039P.

XX 02-OCT-2000; 2000US-237040P.

XX 13-OCT-2000; 2000US-239935P.

XX 20-OCT-2000; 2000US-240960P.

XX 20-OCT-2000; 2000US-241785P.

XX 20-OCT-2000; 2000US-241809P.

XX 01-NOV-2000; 2000US-244617P.

XX 17-NOV-2000; 2000US-249299P.

PR 08-DEC-2000; 2000US-251866P.
PR 08-DEC-2000; 2000US-251866P.
PR 08-DEC-2000; 2000US-251866P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-147444/14.
XX P-PSDB; ABUS5400.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for
PT treating, inhibiting or preventing, e.g. neural, immune system,
PT muscular, respiratory, reproductive, gastrointestinal, pulmonary,
PT cardiovascular or renal disorders -
XX
PS Claim 1; SEQ ID NO 498; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid
CC arthritis and multiple sclerosis), muscular disorders, respiratory
CC diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis),
CC reproductive disorders, gastrointestinal disorders, pulmonary disorders,
CC cardiovascular disorders (e.g. congenital heart defects, Ebstein's
CC anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute
CC kidney failure and end-stage renal disease), hyperproliferative disorders
CC (e.g. Hodgkin's disease and leukemia), inflammatory diseases (e.g.
CC septic shock, bursitis and appendicitis), allergic reactions and
CC conditions (e.g. asthma), blood related disorders (e.g. thrombosis,
CC atherosclerosis and myocardial infarction) and cancerous diseases.
CC Sequences ABX73173-ABX74167 represent human novel polynucleotides of the
CC invention.
XX
SQ Sequence 836 BP; 224 A; 179 C; 221 G; 211 T; 1 other;
XX
Query Match 100.0%; Score 124; DB 25; Length 836;
Best Local Similarity 100.0%; Pred. No. 1e-55;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CGATACGTGGCGGCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 60
DB 209 CGATACGTGGCGGCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 268
XX
QY 61 AAACAACAAGAGAGACTGTGTGTGCTGTGGGAGAGATGTAATCATCTTCCACAACTG 120
DB 269 AAACAACAAGAGAGACTGTGTGTGCTGTGGGAGAGATGTAATCATCTTCCACAACTG 328
XX
QY 121 CTGC 124
DB 329 CTGC 332
XX
RESULT 12
AAC77504
ID AAC77504 standard; cDNA; 962 BP.
XX
AC AAC77504;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF3059 polynucleotide sequence SEQ ID NO:6117.
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
KW anticovulant; osteopathic; antiarthritic; immunosuppressant; cardant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antineoplastic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;

KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
DR WPI, 2000-602362/57.
XX P-PSDB; AAB43295.
XX
DR Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorder,
XX neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 5; Page 5300; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
XX osteopathic; anticovulant; antiarthritic; immunosuppressant;
XX immunostimulant; cardant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antineoplastic;
XX antithyroid; and antianaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorder, asthma,
XX allergic, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 962 BP; 265 A; 204 C; 243 G; 250 T; 0 other;
XX
Query Match 100.0%; Score 124; DB 21; Length 962;
Best Local Similarity 100.0%; Pred. No. 1e-55;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CGATACGTGGCGGCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 60
DB 192 CGATACGTGGCGGCATCTGCAGGGTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 251
XX
QY 61 AAACAACAAGAGAGACTGTGTGTGCTGTGGGAGAGATGTAATCATCTTCCACAACTG 120
DB 252 AAACAACAAGAGAGACTGTGTGTGCTGTGGGAGAGATGTAATCATCTTCCACAACTG 311
XX
QY 121 CTGC 124
DB 312 CTGC 315

RESULT 13
 AAX87324
 ID AAX87324 standard; cDNA; 754 BP.
 XX
 AC AAX87324;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE Human sensitive to apoptosis (SAG) gene mutant MM8.
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KM neurodegenerative disease; muscular dystrophy; wound healing;
 XX vulnerable; therapy; mutant; ds.
 XX Homo sapiens.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH 1..342
 FT CDS /*tag= a
 FT mutation replace(262,T)
 FT /*tag= b
 FT /note= "C88S mutation"
 XX
 PN WO9332514-A2.
 XX
 PD 01-JUL-1999.
 XX
 PF 15-DEC-1998; 98WO-US26705.
 XX
 PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 DR WPI; 1999-430152/36.
 DR P-PSDB; AAY06502.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 XX
 PS Claim 15; Page 71; 84pp; English.
 XX
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM8, which codes for a SAG protein (see AAY06502)
 CC in which the Cys residue at position 88 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 88 from TGC to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-finger 2 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM8
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 XX Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
 SO
 Query Match 97.6%; Score 121; DB 20; Length 754;
 Best Local Similarity 100.0%; Pred. No. 3.9e-54;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGATACGTCGCCATCTGCAGGTCACGATGATGCTGCTTGTAGATGCAAGCTGA 60
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 141 CGATACGTCGCCATCTGCAGGTCACGATGATGCTGCTTGTAGATGCAAGCTGA 200
 QY 61 AACACAACAGAGGACTGTGTGTGCTGGGAGATGAATCATCTCCACACATG 120
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 201 AACACAACAGAGGACTGTGTGTGCTGGGAGATGAATCATCTCCACACATG 260
 QY 121 C 121
 Db 261 C 261
 RESULT 14
 AAX87317
 ID AAX87317 standard; cDNA; 754 BP.
 XX
 AC AAX87317;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE Human sensitive to apoptosis (SAG) gene mutant MM1.
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KM neurodegenerative disease; muscular dystrophy; wound healing;
 XX vulnerable; therapy; mutant; ds.
 XX Homo sapiens.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH 1..342
 FT CDS /*tag= a
 FT mutation replace(148,T)
 FT /*tag= b
 FT /note= "C50S mutation"
 XX
 PN WO9332514-A2.
 XX
 PD 01-JUL-1999.
 XX
 PF 15-DEC-1998; 98WO-US26705.
 XX
 PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 DR WPI; 1999-430152/36.
 DR P-PSDB; AAY06495.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 XX
 PS Claim 15; Page 59; 84pp; English.
 XX
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM1, which codes for a SAG protein (see AAY06495)
 CC in which the Cys residue at position 50 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 50 from TGC to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is at a haem binding site of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. MM1 showed reduced haem binding but unaffected
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC


```

XX 01-JUL-1999.
PD 15-DEC-1998; 98MO-US26705.
XX
PF 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.
XX
PA (WARN ) WARNER LAMBERT CO.
XX
PI Sun Y;
XX
DR WPI; 1999-430152/36.
DR P-PSDB; AAY06499.
XX
PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
XX
PS Claim 15; Page 66; 84pp; English.
XX
CC This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM5, which codes for a SAG protein (see AAY06499)
CC in which the Cys residue at position 80 of the native protein (see
CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
CC 80 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is in zinc-finger finger 1 of SAG. Single and
CC double SAG mutants (see AAX87317-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM5
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversion of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.
XX
SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 78.2%; Score 97; DB 20; Length 754;
Best Local Similarity 100.0%; Pred. No. 1.9e-41;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATAGCTGGCCATCTGCAGGCTCCAGTGATGATGCTGTCTTAGATGTAAGCTGA 60
DB 141 CGATAGCTGGCCATCTGCAGGCTCCAGTGATGATGCTGTCTTAGATGTAAGCTGA 200
QY 61 AACCAACAAGAGAGACTGTGTGTGCTGGGAGAA 97
DB 201 AACCAACAAGAGAGACTGTGTGTGCTGGGAGAA 237

RESULT 21
ID AAX87315 standard; cDNA; 747 BP.
XX
AC AAX87315;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene mutant 1.
XX
KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
KW neurodegenerative disease; muscular dystrophy; wound healing;
KW vulnereary; therapy; mutant; colon carcinoma; ds.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT 1..273
FT CDS
FT /tag= a
XX

```

```

PN WO9932514-A2.
PD 01-JUL-1999.
XX
PF 15-DEC-1998; 98MO-US26705.
XX
PR 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.
XX
PA (WARN ) WARNER LAMBERT CO.
XX
PI Sun Y;
XX
DR WPI; 1999-430152/36.
DR P-PSDB; AAY06493.
XX
PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
XX
PS Claim 15; Page 54-55; 84pp; English.
XX
CC This is the nucleotide sequence of deletion mutant 1 of the novel
CC human sensitive to apoptosis gene SAG (see also AAX87314). The cDNA
CC was obtained by PCR amplification of RNA isolated from DLD-1 colon
CC carcinoma cells (ATCC CCL 221). It contains a 7 bp deletion.
CC compared with native SAG, starting at nucleotide 170 of the coding
CC region that codes for a potential haem binding site. The frame
CC shift deletion abolishes the downstream zinc finger in the
CC resulting protein (see AAY06493). The mutation was detected by PCR
CC in SAG RNA, but not in genomic DNA. It suggests a possible role
CC for SAG in human carcinogenesis. SAG genes, and mutant SAG
CC genes, can be used to protect cells from apoptosis induced by redox
CC reagents. Antisense SAG genes can be used to inhibit the growth of
CC tumour cells. The SAG genes can also be used for the recombinant
CC production of the SAG proteins. The SAG proteins can be used to
CC scavenge oxygen radicals in organisms and to promote wound healing.
CC Additionally, the SAG genes or their complements can be used to
CC promote or inhibit the growth of plant cells (all claimed). The SAG
CC protein is also an ideal molecular target in the development of
CC drugs against neurodegenerative disorders, cancers and muscle
CC dystrophy.
XX
SQ Sequence 747 BP; 204 A; 155 C; 197 G; 191 T; 0 other;

Query Match 72.6%; Score 90; DB 20; Length 747;
Best Local Similarity 100.0%; Pred. No. 9.7e-38;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GATGCTGTCTTAGATGTAAGCTGAACCAACAAGAGAGACTGTGTGCTGGGGA 94
DB 168 GATGCTGTCTTAGATGTAAGCTGAACCAACAAGAGAGACTGTGTGCTGGGGA 227
QY 95 GATGTAATCATCTCTCCACAACACTGCTGC 124
DB 228 GATGTAATCATCTCTCCACAACACTGCTGC 257

RESULT 22
ID AAH25847 standard; DNA; 1152 BP.
XX
AC AAH25847;
XX
DT 20-AUG-2001 (first entry)
XX
DE Human apoptosis associated protein 12 coding sequence.
XX
KW Human; apoptosis associated protein 12; SAG protein 12; cancer;
KW haemopathy; HIV infection; immunological disease; inflammation; ds.
XX
OS Homo sapiens.
XX
PH WO200132863-A1.
XX

```


XX PD 10-MAY-2001.
XX PA 30-OCT-2000; 2000WO-CN00406.
XX PF 29-OCT-1999; 99CN-0119924.
XX PR (SHAN-) SHANGHAI BIO DOOR GENE TECHNOLOGY LTD.
XX PA Mao Y, Xie Y;
XX PI WPI; 2001-335831/35.
XX DR P-PSDB; AAB98975.
XX PT Human apoptosis associated protein 12 and encoded polynucleotide,
XX PT applicable in diagnosis and treatment of malignant tumour, hemopathy,
XX PT HIV infection, immunological diseases and various inflammation -
XX PS Claim 6; Page 22-23; 27pp; Chinese.
XX CC The present invention provides the protein and coding sequences of human
XX CC apoptosis associated (SAG) protein 12. These sequences can be used in
XX CC the diagnosis and treatment of malignant tumours, haemopathy, HIV
XX CC infection, immunological diseases and various types of inflammation. The
XX CC present sequence is the SAG protein 12 coding sequence.
SQ Sequence 1152 BP; 292 A; 265 C; 309 G; 286 T; 0 other;
Query Match 71.8%; Score 89; DB 22; Length 1152;
Best Local Similarity 100.0%; Pred. No. 3.2e-37;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 36 ATGCCCTGTTAGATGTCAGTGAAGTGAACAAACAGAGACTGTGTGCTGGGAG 95
DB 564 ATGCCCTGTTAGATGTCAGTGAAGTGAACAAACAGAGACTGTGTGCTGGGAG 623
QY 96 AATGTATCATTCCTCCACAACTGCTGC 124
DB 624 AATGTATCATTCCTCCACAACTGCTGC 652
RESULT 23
AAK87320
ID AAK87320 standard; cDNA; 754 BP.
XX AC AAK87320;
XX DT 27-SEP-1999 (first entry)
XX DE Human sensitive to apoptosis (SAG) gene mutant MM4.
XX KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX KW neurodegenerative disease; muscular dystrophy; wound healing;
XX KW vulnerability; therapy; mutant; ds.
XX OS Homo sapiens.
XX OS Synthetic.
FH Key Location/Qualifiers
FT CDS 1..342 /*tag= a
FT mutation replace(181,T)
FT /*tag= b
FT /note= "C61S mutation"
XX PN WO9932514-A2.
XX PD 01-JUL-1999.
XX PF 15-DEC-1998; 98WO-US26705.
XX PR 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.

XX PA (WARN) WARNER LAMBERT CO.
XX PI Sun Y;
XX PI WPI; 1999-430152/36.
XX DR P-PSDB; AAY06498.
XX PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
XX PT promoting cell growth and protecting cells against apoptosis
XX PS Claim 15; Page 64-65; 84pp; English.
XX CC This is the nucleotide sequence of human sensitive to apoptosis
XX CC (SAG) mutant gene MM4, which codes for a SAG protein (see AAY06498)
XX CC in which the Cys residue at position 61 of the native protein (see
XX CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
XX CC 50 from TGC to AGC obtained by site-directed mutagenesis of SAG
XX CC cDNA. This residue is at zinc-finger finger 1 of SAG. Single and
XX CC double SAG mutants (see AAK87317-31) were made in order to determine
XX CC the role of each cysteine residue of SAG in haem binding and SAG
XX CC oligomerization. These properties were unaffected by the MM4
XX CC mutation. SAG is a novel zinc finger protein that promotes
XX CC cell growth, protects cells from apoptosis, scavenges oxygen
XX CC radicals and can be used for the reversion of a tumour phenotype.
XX CC SAG genes, and mutant SAG genes, can be used to protect cells from
XX CC apoptosis induced by redox reagents. They can also be used for the
XX CC recombinant production of SAG proteins, which are molecular targets
XX CC in the development of drugs against neurodegenerative disorders,
XX CC cancers and muscle dystrophy, and promoting wound healing.
SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
Query Match 66.9%; Score 83; DB 20; Length 754;
Best Local Similarity 100.0%; Pred. No. 4.9e-34;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 GTCTTAGATGTCAGTGAAGTGAACAAACAGAGACTGTGTGCTGGGAGATGTA 101
DB 182 GTCTTAGATGTCAGTGAAGTGAACAAACAGAGACTGTGTGCTGGGAGATGTA 241
QY 102 ATCAATCCTTCACAACTGCTGC 124
DB 242 ATCAATCCTTCACAACTGCTGC 264
RESULT 24
AAK87328
ID AAK87328 standard; cDNA; 754 BP.
XX AC AAK87328;
XX DT 27-SEP-1999 (first entry)
XX DE Human sensitive to apoptosis (SAG) gene mutant MM12.
XX KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX KW neurodegenerative disease; muscular dystrophy; wound healing;
XX KW vulnerability; therapy; mutant; ds.
XX OS Homo sapiens.
XX OS Synthetic.
FH Key Location/Qualifiers
FT CDS 1..342 /*tag= a
FT mutation replace(217,T)
FT /*tag= b
FT /note= "C73S mutation"
XX PN WO9932514-A2.
XX PD 01-JUL-1999.

XX 15-DEC-1998; 98WO-US26705.
 PF 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX (WARN) WARNER LAMBERT CO.
 PA
 XX
 PI Sun Y;
 DR WPI, 1999-430152/36.
 DR P-PSDB; AAY06506.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PS
 PS Claim 15; Page 77-78; 84pp; English.
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM12, which codes for a SAG protein (see AAY06506)
 CC in which the Cys residue at position 73 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 73 from TGT to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in a proase inhibitor motif of SAG. Single
 CC and double SAG mutants (see AAX87317-31) were made to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM12
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
 Query Match 61.3%; Score 76; DB 20; Length 754;
 Best Local Similarity 100.0%; Pred. No. 2.4e-30;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGATAGCTGCGCCATCTGAGGTCGAGTGATGATGCTGCTTAGATGTAAGCTGA 60
 DB 141 CGATAGCTGCGCCATCTGAGGTCGAGTGATGATGCTGCTTAGATGTAAGCTGA 200
 QY 61 AAACAACAAGAGGAC 76
 DB 201 AAACAACAAGAGGAC 216
 RESULT 25
 AAX87327
 ID AAX87327 standard; cDNA; 754 BP.
 XX
 AC AAX87327;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE Human sensitive to apoptosis (SAG) gene mutant MM11.
 XX
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnerability; therapy; mutant; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FT Key Location/Qualifiers.
 FT CDS 1..342
 FT mutation /tag= a
 FT /tag= b

FT /note= "C64S mutation"
 XX
 PN WO932514-A2.
 XX
 PD 01-JUL-1999.
 XX
 PF 15-DEC-1998; 98WO-US26705.
 XX
 PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 DR WPI, 1999-430152/36.
 DR P-PSDB; AAY06505.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PS
 PS Claim 15; Page 76; 84pp; English.
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM11, which codes for a SAG protein (see AAY06505)
 CC in which the Cys residue at position 64 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 64 from TGT to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-finger 1 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM11
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
 Query Match 59.7%; Score 74; DB 20; Length 754;
 Best Local Similarity 100.0%; Pred. No. 2.8e-29;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 GTCAAGCTGAACAACAAGAGAGACTGTGTTGTTGGTGGGAGAGATTAATCAATTCCT 110
 DB 191 GTCAAGCTGAACAACAACAAGAGAGACTGTGTTGTTGGTGGGAGAGATTAATCAATTCCT 250
 QY 111 TCCACAACTGCTGC 124
 DB 251 TCCACAACTGCTGC 264
 RESULT 26
 AAX87329
 ID AAX87329 standard; cDNA; 754 BP.
 XX
 AC AAX87329;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE Human sensitive to apoptosis (SAG) gene mutant MM13.
 XX
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnerability; therapy; mutant; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS

[illegible]

DE	Human secreted protein 5' EST SEQ ID NO: 297.
XX	
KW	Human; secreted protein; EST; expressed sequence tag; diagnosis;
KM	forensic; gene therapy; chromosome mapping; signal peptide;
KW	upstream regulatory sequence; cytokine activity; cell proliferation;
KM	differentiation; haematopoiesis regulation; tissue growth regulation;
KX	reproductive hormone regulation; chemotactic; chemokine; haemostatic;
XX	thrombolytic; anti-inflammatory; tumour inhibition; ds.
OS	Homo sapiens.
XX	
PN	WO906554-A2.
XX	
PD	11-FEB-1999.
XX	
PJ	31-JUL-1998; 98WO-IB01238.
PF	
PR	01-AUG-1997; 97US-0905134.
XX	
PA	(GEST) GENSET.
PI	Duclet A, Dumas Milne Edwards J, Lacroix B;
DR	WPt, 1999-153784/13.
DR	P-PsDB; AAY12252.
XX	
PT	New nucleic acids encoding human secreted proteins - obtained from
PT	cDNA libraries prepared from kidney, fetal kidney, dystrophic
PT	muscle, muscle and heart tissue
PS	Claim 1; Page 441; 622pp; English.
XX	
CC	AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
CC	human secreted proteins, and encode the proteins given in AAY01602 and
CC	AAY11994 to AAY12260, respectively. The proteins given represent the
CC	signal peptide and an N-terminal fragment of a secreted protein. The
CC	nucleic acid sequences can be used for producing secreted human gene
CC	products. They can also be used to develop products for diagnosis and
CC	therapy. The proteins obtained may have cytokine activity, cell
CC	proliferation/differentiation activity, haematopoiesis regulating
CC	activity, tissue growth regulating activity, reproductive hormone
CC	regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC	thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC	activity, tumour inhibition activity or other activities. The products
CC	can be used in forensic, gene therapy and chromosome mapping procedures.
CC	The sequences can also be used for obtaining corresponding promoter
CC	sequences. The nucleic acids encoding the signal peptide can be used
CC	for directing extracellular secretion of a polypeptide or the insertion
CC	of a polypeptide into a membrane, or importing a polypeptide into
CC	a cell.
SQ	
Sequence	224 BP; 41 A; 60 C; 83 G; 38 T; 2 other;
Query Match	34.7%; Score 43; DB 20; Length 224;
Best Local Similarity	100.0%; Pred. No. 7e-13;
Matches	43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 CGATACGTGGGCCCATCTGCAGGGTTCGAGTGATGCCTGT 43 Dd 164 CGATACGTGGGCCCATCTGCAGGGTTCGAGTGATGCCTGT 206
RESULT 28	
ID	AAX87316
XX	AAX87316 standard; cDNA; 706 BP.
AC	
XX	AAX87316;
DT	27-SEP-1999 (first entry)
DE	Human sensitive to apoptosis (SAG) gene mutant 2.
XX	
SAG	gene; sensitive to apoptosis; human; cancer; tumour;
XX	

RESULT 30
ABN40537
ID ABN40537 standard; DNA; 60 BP.
XX
AC ABN40537;
XX
XX 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:13285.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-1B01903.
XX
XX 28-JUL-2000; 2000US-221607P.
XX
XX 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Masserman A, Mintz E, Mintz L, Faigler S;
XX
XX MPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
XX
PS Example 1; SEQ ID 13285; 47bp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridizing selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterizing the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the
XX detection of tissue- and pathology-specific genes such as those genes
XX only expressed in specific tissue under a specific pathological
XX condition; to detect developmental specific genes; and to detect RNA
XX transcripts and splice variants of a transcriptome of a patient suffering
XX from a particular disorder. ABN27253 to ABN55589 represent
XX oligonucleotide sequences from rats, humans and mice, which are used in
XX the exemplification of the present invention.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 10 A; 13 C; 25 G; 12 T; 0 other;
XX
Query Match 28.2%; Score 35; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTGGCCATCTGCAGGGTCCAGGTATGG 35
DB 26 CGATACGTGGCCATCTGCAGGGTCCAGGTATGG 60

RESULT 31
ABZ11414
ID ABZ11414 standard; CDNA; 439 BP.
XX
XX
AC ABZ11414;
XX
XX 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 296.
XX
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
XX cell-proliferative disorder; neurodegenerative disease; bacterial;
XX Parkinson's disease; Alzheimer's disease; autoimmune disease;
XX multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
XX arthritis; cytostatic; immunomodulator; neurotropic; neuroprotective;
XX antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
XX haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
XX antiarthritic; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200270539-A2.
XX
XX 12-SEP-2002.
XX
XX 05-MAR-2002; 2002WO-US05095.
XX
XX 05-MAR-2001; 2001US-0799451.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YF, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
XX Wehman T, Wang J, Wang D, Drmanac RT;
XX
XX MPI; 2002-759812/82.
XX
XX P-PSDB; ABP69197.
XX
XX New polynucleotides comprising sequences assembled from expressed
XX sequence tags (ESTs), useful for treating cell-proliferative,
XX neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
XX platelet or coagulation disorders -
XX
XX
PS Claim 1; SEQ ID NO 296; 1012pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising a
XX nucleotide sequence selected from any of 948 sequences
XX (ABZ1119-ABZ12066) or their mature protein coding portion, active domain
XX (ABZ1119-ABZ12066) or complementary sequences. The polynucleotides are useful
XX for identifying expressed genes or for physical mapping of human genome.
XX The encoded polypeptides (ABP68902-ABP69843) are useful as molecular
XX weight markers, as a food supplement, for generating antibodies, in
XX medical imaging, screening and diagnostic assays and for treating
XX cell-proliferative disorders (cancer), neurodegenerative diseases
XX (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
XX sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
XX disorders, platelet or coagulation disorders, wound, burns, infection,
XX ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
XX parasitic), arthritis, etc.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 439 BP; 69 A; 142 C; 162 G; 66 T; 0 other;
XX

QY 1 CGATACGTGGCCATCTGCAGGGTCCAGGTATGG 35
DB 221 CGATACGTGGCCATCTGCAGGGTCCAGGTATGG 255

XX	RESULT 32	
XX	AAx87313	
XX	ID	AAx87313 standard, cDNA, 1140 BP.
XX	AC	AAx87313;
XX	DT	27-SEP-1999 (first entry)
XX	DE	Mouse sensitive to apoptosis (SAG) gene.
XX	KW	SAG gene; sensitive to apoptosis; mouse; cancer; tumour;
XX	KW	neurodegenerative disease; muscular dystrophy; wound healing;
XX	KW	vulnerary; therapy; ds.
XX	OS	Mus musculus.
XX	FT	Key
XX	FT	Location/Qualifiers
XX	FT	CDS
XX	FT	17..358
XX	FT	/*tag= a
XX	PN	W09932514-A2.
XX	PD	01-JUL-1999.
XX	PF	15-DEC-1998; 98WO-US26705.
XX	PR	11-SEP-1998; 98US-0099840.
XX	PR	19-DEC-1997; 97US-0068179.
XX	PA	(WARN) WARNER LAMBERT CO.
XX	PI	Sun Y;
XX	DR	WPI; 1999-430152/36.
XX	DR	P-PSDB; AAY06491.
XX	PT	SAG: Sensitive to Apoptosis Gene and related proteins, useful for
XX	PT	promoting cell growth and protecting cells against apoptosis
XX	PS	Claim 1; Page 48-49; 84pp; English.
XX	PS	This is the nucleotide sequence of mouse sensitive to apoptosis
XX	PS	(SAG) cDNA, which codes for a novel redox-sensitive, haem-binding
XX	PS	protein (see AAY06491) with a zinc RING finger domain that promotes
XX	PS	cell growth, protects cells from apoptosis, scavenges oxygen
XX	PS	radicals and can be used for the reversion of a tumour phenotype.
XX	PS	The cDNA was isolated using differential display to identify genes
XX	PS	associated with 1,10-phenanthroline (OP)-induced apoptosis in
XX	PS	murine tumour lines, and use of an isolated OP-inducible clone to
XX	PS	screen a mouse lung cDNA library. The mouse SAG cDNA was used to
XX	PS	identify human SAG (see AAX87314). SAG is highly conserved among
XX	PS	species. Disruption in yeast was shown to be lethal. SAG deletion
XX	PS	mutants (see AAX87315-16) have been identified in human cancer lines,
XX	PS	suggesting a role in carcinogenesis. SAG genes, and mutant SAG
XX	PS	genes, can be used to protect cells from apoptosis induced by redox
XX	PS	reagents. Antisense SAG genes can be used to inhibit the growth of
XX	PS	tumour cells. The SAG genes can also be used for the recombinant
XX	PS	production of the SAG proteins. The SAG proteins can be used to
XX	PS	scavenge oxygen radicals in organisms and to promote wound healing.
XX	PS	Additionally, the SAG genes or their complements can be used to
XX	PS	promote or inhibit the growth of plant cells (all claimed). The SAG
XX	PS	protein is also an ideal molecular target in the development of
XX	PS	drugs against neurodegenerative disorders, cancers and muscle
XX	PS	dystrophy.
XX	Sequence	1140 BP; 302 A; 224 C; 287 G; 327 T; 0 other;
XX	Query Match	25.8%; Score 32; DB 20; Length 1140;
XX	Best Local Similarity	100.0%; Pred. No. 4.5e-07;
XX	Matches	32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	11	GCCATCTGCAGAGGTCACAGTGATGATGCCCTG 42

```

Db      167 GCCATCTGCAGGCGTCCACAGTGATGGATGCTTC   198
|||||
RESULT 33
AL16209 standard; cDNA; 264 BP.
XX AL16209,
XX
XX 07-DEC-2001 (first entry)
XX DE Human breast cancer expressed polynucleotide 8666.
XX KW Human; breast cancer; cell marker; cytostatic; ss.
XX OS Homo sapiens.
XX PN W0200151628-A2.
XX PD
XX PF 10-JAN-2001; 2001WO-US00798.
XX PR 14-JAN-2000; 2000US-0176077.
XX PR 14-MAR-2000; 2000US-0189167.
XX PR 24-MAR-2000; 2000US-0192099.
XX PR 29-MAR-2000; 2000US-0193480.
XX PR 15-MAY-2000; 2000US-0205230.
XX PR 09-JUN-2000; 2000US-0211315.
XX PR 25-JUL-2000; 2000US-0220534.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX DR WPI; 2001-451856/48.
XX PT New peptide useful as a marker for the diagnosis of breast cancer -
XX PS Claim 1; Page 1568; 3695pp; English.
XX CC The invention relates to human breast cancer expressed polynucleotides
XX CC (AAU07544-AAU26789) and methods of assessing whether a patient is
XX CC afflicted with breast cancer by examining the correlation between the
XX CC expression of certain markers and the cancerous state of breast cells.
XX CC The polynucleotides and encoded polypeptides are potential markers for
XX CC detecting, diagnosing, monitoring, characterising treating and
XX CC potentially preventing breast cancer. The polynucleotides and encoded
XX CC polypeptides are also useful for isolating compounds with cytostatic
XX SO activity.
XX Sequence 264 BP; 64 A; 59 C; 74 G; 62 T; 5 other;
OY Query Match 19.4%; Score 24; DB 22; Length 264;
DB Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 101 AAATCATCCTTCACCAACTGTCTGC 124
DB 54 AAATCATCCTTCACCAACTGTCTGC 77
RESULTS 34
ID AL162052 standard; cDNA; 596 BP.
XX AL162052,
XX AC
XX AL162052;
XX AC
XX 07-DEC-2001 (first entry)
XX DT
XX DE Human breast cancer expressed polynucleotide 17509.
XX
```

```
KM Human; breast cancer; cell marker; cyrostatic; ss.
XX
XX Homo sapiens.
OS
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US00798.
XX
XX 14-JAN-2000; 2000US-0176077.
XX
XX 14-MAR-2000; 2000US-0189167.
XX
XX 24-MAR-2000; 2000US-0192099.
XX
XX 29-MAR-2000; 2000US-0193480.
XX
XX 15-MAY-2000; 2000US-0205230.
XX
XX 09-JUN-2000; 2000US-0211315.
XX
XX 25-JUL-2000; 2000US-0220534.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer -
XX
XX Claim 1; Page 3235; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
XX (AA107544-AA126789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterizing treating and
XX potentially preventing breast cancer. The polynucleotides and encoded
XX polypeptides are also useful for isolating compounds with cyrostatic
XX activity.
XX
XX Sequence 596 BP; 166 A; 124 C; 144 G; 162 T; 0 other;
XX
XX Query Match 19.4%; Score 24; DB 22; Length 596;
XX Best Local Similarity 100.0%; Pred. No. 0.0077;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 101 AATCATTCCTTCACAACTGCTGC 124
XX |||||||||||||||||||
XX Db 11 AATCATTCCTTCACAACTGCTGC 34
XX
XX RESULT 35
XX AAL22239
XX ID AAL22239 standard; cDNA; 630 BP.
XX
XX AAL22239;
XX
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 14696.
XX
XX Human; breast cancer; cell marker; cyrostatic; ss.
XX
XX Homo sapiens.
XX
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US00798.
XX
XX 14-JAN-2000; 2000US-0176077.
XX
XX 14-MAR-2000; 2000US-0189167.
XX
XX 24-MAR-2000; 2000US-0192099.
XX
XX 29-MAR-2000; 2000US-0193480.
XX
XX 15-MAY-2000; 2000US-0205230.
XX
XX 09-JUN-2000; 2000US-0211315.
XX
XX 25-JUL-2000; 2000US-0220534.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer -
XX
XX Claim 1; Page 3235; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
XX (AA107544-AA126789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterizing treating and
XX potentially preventing breast cancer. The polynucleotides and encoded
XX polypeptides are also useful for isolating compounds with cyrostatic
XX activity.
XX
XX Sequence 596 BP; 166 A; 124 C; 144 G; 162 T; 0 other;
XX
XX Query Match 19.4%; Score 24; DB 22; Length 596;
XX Best Local Similarity 100.0%; Pred. No. 0.0077;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer -
XX
XX Claim 1; Page 2649-2650; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
XX (AA107544-AA126789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterizing treating and
XX potentially preventing breast cancer. The polynucleotides and encoded
XX polypeptides are also useful for isolating compounds with cyrostatic
XX activity.
XX
XX Sequence 630 BP; 213 A; 125 C; 121 G; 171 T; 0 other;
XX
XX Query Match 19.4%; Score 24; DB 22; Length 630;
XX Best Local Similarity 100.0%; Pred. No. 0.0077;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 101 AATCATTCCTTCACAACTGCTGC 124
XX |||||||||||||||||||
XX Db 43 AATCATTCCTTCACAACTGCTGC 66
XX
XX RESULT 36
XX AAA70390/C
XX ID AAA70390 standard; DNA; 45 BP.
XX
XX AAA70390;
XX
XX 02-FEB-2001 (first entry)
XX
XX Site directed mutagenesis oligonucleotide # 1 for placental bikunin.
XX
XX Mucociliary dysfunction; mucus; sputum; human;
XX chronic obstructive lung disease; chronic bronchitis; CB; Bronchiectasis;
XX BE; asthma; cystic fibrosis; CF; bacterial infection; placental bikunin;
XX Kunitz-type serine protease inhibitor; chronic sinusitis; glue ear; ss.
XX
XX Homo sapiens.
XX
XX WO200037099-A2.
XX
XX 29-JUN-2000.
XX
XX 22-DEC-1999; 99WO-GB04381.
XX
XX 22-DEC-1998; 98US-0218913.
XX
XX 17-NOV-1999; 99US-0441966.
XX
XX (FARB ) BAYER AG.
XX
XX Hall R, Poll CT, Newton BB, Taylor WJA;
XX
XX WPI; 2000-452127/39.
XX
XX Stimulating mucociliary clearance rate of mucus and sputum in lung
XX airways for treating lung diseases such as cystic fibrosis and
XX bronchitis involves administering a Kunitz-type serine protease
XX inhibitor -
XX
XX Example 9; Page 59; 173pp; English.
```

XX Mucociliary dysfunction is the inability of ciliated epithelium to clear
 CC mucus and sputum in lung airways. Mucociliary dysfunction is a serious
 CC complication of chronic obstructive lung diseases such as Chronic
 CC Bronchitis (CB), Bronchiectasis (BE), asthma and Cystic Fibrosis (CF).
 CC In addition, patients suffering from mucociliary dysfunction are
 CC susceptible to secondary bacterial infections. A sequence for human
 CC placental bikunin has been isolated (see AAB14190 for the protein
 CC sequence). Placental bikunin is a Kunitz-type serine protease inhibitor
 CC and protein, which can stimulate the rate of mucociliary clearance of mucus
 CC and sputum in lung airways. Therefore, placental bikunin protein may be
 CC used for treating lung diseases such as CF, CB, BE, and chronic
 CC sinusitis and glue ear which are caused by retention and accumulation of
 CC mucus. The present sequence is an oligonucleotide used in site-directed
 CC mutagenesis of the coding sequence of the human placental bikunin protein
 CC sequence of AAB14190.

XX
 XX Sequence 45 BP; 4 A; 15 C; 18 G; 8 T; 0 other;
 SQ

Query Match 15.3%; Score 19; DB 21; Length 45;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 ACGTGGCCCATCTGCAGG 23
 |||||
 38 ACGTGGCCCATCTGCAGG 20

Db

RESULT 37
 AAX87337
 ID AAX87337 standard; DNA; 18 BP.
 XX
 AC AAX87337;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE Human sensitive to apoptosis gene (SAG) primer hSAG.M1.
 XX
 KW SAG gene: sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnerability; therapy; PCR; primer; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN MO932514-A2.
 XX
 PD 01-JUL-1999.
 XX
 PF 15-DEC-1998; 98MO-US26705.
 XX
 PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 XX
 DR WPI; 1999-430152/36.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 XX
 PS Example 13, Page 26; 84pp; English.
 XX
 CC Primer hSAG.M1 corresponds to nucleotides 151-171 of the human
 CC sensitive to apoptosis gene (SAG) cDNA clone provided in AAX87314.
 CC It was used with primer SAGT-02-1 (see AAX87338) in the RT-PCR
 CC amplification of RNA isolated from 20 human tumour lines and
 CC transformed lines originating from lung, brain, kidney, prostate,
 CC testis, nasopharynx, bone, cervix and foreskin. 2 SAG deletion
 CC mutants (see AAX87315-16) were detected in cancer cell lines
 CC originating from colon and testis, suggesting a possible role for

CC SAG in human carcinogenesis. SAG (see also AAY06492) is a
 CC redox-sensitive, haem-binding protein that promotes cell growth.
 CC protects cells from apoptosis, and scavenges oxygen radicals. It
 CC can be used to reverse a tumour phenotype.
 XX

SQ Sequence 18 BP; 3 A; 6 C; 6 G; 3 T; 0 other;
 SQ

Query Match 14.5%; Score 18; DB 20; Length 18;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 GCCATCTGCAGGTCAG 28
 |||||
 1 GCCATCTGCAGGTCAG 18

Db

RESULT 38
 AAC29980/c
 ID AAC29980 standard; cDNA; 349 BP.
 XX
 AC AAC29980;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 34055.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 OS
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (BEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 1; SEQ ID 34055; 71pp + CD-ROM; English.
 XX
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC cDNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX

SQ Sequence 349 BP; 101 A; 65 C; 67 G; 116 T; 0 other;
 SQ

Query Match 14.5%; Score 18; DB 21; Length 349;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 GGGAGATGATATCATTC 108

Db 85 GGGAGATGTATCATTC 68

RESULT 39
ABV87139/c
ID ABV87139 standard; cDNA; 541 BP.
XX
XX
AC ABV87139;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human colon cancer related cDNA SEQ ID NO 450.
XX
KW Human; colon; cancer; cytostatic; tumour; gene therapy; vaccine; gene;
XX ss.
XX Homo sapiens.
XX OS
XX PN MO200258534-A2.
XX PD 01-AUG-2002.
XX PF 19-NOV-2001; 2001WO-US43704.
XX PR 20-NOV-2000; 2000US-252222P.
XX PR 06-FEB-2001; 2001US-267011P.
XX PR 28-MAR-2001; 2001US-279670P.
XX PR 10-JUL-2001; 2001US-304037P.
XX PA (CORI-) CORIXA CORP.
XX PI Stolk JA, Xu J, Chenault RA, Meagher MJ, Secrist H, King GE;
XX WPI; 2002-608400/65.
XX DR
XX PT New isolated tumor colon polynucleotide and polypeptide, useful for the
XX PT diagnosis, prevention and/or treatment of cancer, in particular colon
XX PT cancer -
XX PS Claim 1; SEQ ID NO 450; 266pp + Sequence Listing; English.
XX
XX CC The invention relates to a human colon tumour expressed polynucleotide
XX CC (I) encoding a polypeptide (II, ABP67991-ABP67996) comprising: (i) any of
XX CC 2600 fully defined nucleotide sequences (ABV8669-ABV89289); (ii)
XX CC complements of (i); (iii) at least 20 contiguous residues of (i); (iv)
XX CC sequences that hybridize to (i); under moderately stringent conditions;
XX CC (v) sequences having at least 75% or 90% identity to (i); or (vi)
XX CC degenerate variants of (i). The compositions and methods of the present
XX CC invention are useful for the diagnosis, prevention and/or treatment of
XX CC cancer, particularly colon cancer. (I) can be used in gene therapy and
XX CC (i) and (ii) are useful in pharmaceutical compositions such as vaccines.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 541 BP; 173 A; 99 C; 105 G; 164 T; 0 other;
XX
XX Query Match 14.5%; Score 18; DB 24; Length 541;
XX Best Local Similarity 100.0%; Pred. No. 11;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 91 GGGAGATGTATCATTC 108
XX DB 154 GGGAGATGTATCATTC 137

DT 25-MAR-2003 (updated)
DT 17-MAR-1995 (first entry)
XX
XX Protein kinase (CKIgamma1Hu) coding sequence.
XX DE
XX KW Protein kinase; immunogen; antibody; protein-tyrosine kinase;
XX KW protein-serine/threonine kinase; recombination; repair; screening;
XX KW detection; casein kinase; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT 67..1200
XX FT /*tag= a
XX FT /product= Protein kinase.
XX
XX PN M09417189-A2.
XX PD 04-AUG-1994.
XX PF 21-JAN-1994; 94WO-US00795.
XX PR 21-JAN-1993; 93US-0008001.
XX
XX PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX PI Hoeckstra MF;
XX DR WPI; 1994-264102/32.
XX DR P-PSDB; AAR56529.
XX
XX PT Wild-type and mutant protein kinase genes and encoded
XX PT polypeptide(s) - Useful in screening for compositions which may
XX PT effect DNA double strand break repair activity
XX PS Claim 10; Page 97-99; 121pp; English.
XX
XX CC This sequence encodes a protein kinase which is an HRR25-like
XX CC protein. Host cells stably transformed with the protein kinase
XX CC encoding DNA may be used for the expression of the protein kinase
XX CC such that the expressed protein is "displayed" on the host cell
XX CC surface. The cells may then be used as immunogen for the production
XX CC of antibodies. The host cells may also be used for the large scale
XX CC production of the protein kinase. The expressed protein being either
XX CC isolated from the cell surface or from the culture medium.
XX CC Recombinant HRR25 like proteins display a number of properties which
XX CC are unique among the eukaryotic protein kinases e.g. the HRR25
XX CC protein possesses both protein-tyrosine kinase and
XX CC protein-serine/threonine kinase activities. Also, HRR25 operates to
XX CC promote repair of DNA strand breaks at a specific nucleotide sequence
XX CC and is the only protein kinase known to have such
XX CC recombination/repair promoting activity. Recombinant HRR25-like
XX CC proteins and host cells expressing them are useful in screening
XX CC methods designed to examine the effects of various compositions on
XX CC DNA break repair and protein kinase activities of the protein. The
XX CC HRR25-like proteins are casein kinase class I protein kinases.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 2405 BP; 849 A; 392 C; 486 G; 678 T; 0 other;
XX
XX Query Match 14.5%; Score 18; DB 15; Length 2405;
XX Best Local Similarity 100.0%; Pred. No. 11;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 91 GGGAGATGTATCATTC 108
XX DB 2241 GGGAGATGTATCATTC 2224

RESULT 40
AAQ70827/c
ID AAQ70827 standard; cDNA; 2405 BP.
XX
XX AC AAQ70827;
XX

RESULT 41
AAQ92962/c
ID AAQ92962 standard; DNA; 2405 BP.
XX

XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI MPI; 2001-639362/73.
DR P-PSDB; ABG20443.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 20434; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. A564197-A564564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 239 BP; 47 A; 87 C; 58 G; 47 T; 0 other;
Query Match 13.7%; Score 17; DB 23; Length 239;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 CCATCTGCAGGTCACG 28
Db 139 CCATCTGCAGGTCACG 123
RESULT 44
ABA46329
ID ABA46329 standard; DNA; 355 BP.
XX
XX ABA46329;
AC
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #5024.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX
OS Homo sapiens.
XX
FN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR

PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0633366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI MPI; 2001-496933/54.
XX
XX
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes
XX
PS Claim 1; SEQ ID NO 5024; 327pp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;
Query Match 13.7%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22 GGTCCAGGTGATGATG 38
Db 281 GGTCCAGGTGATGATG 297
RESULT 45
ABA51431
ID ABA51431 standard; DNA; 355 BP.
XX
XX ABA51431;
AC
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #10126.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX
OS Homo sapiens.
XX
FN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR

PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes -
PS Claim 4; SEQ ID NO 10126; 327bp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;
XX
QY Query Match 13.7%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 22 GGTCAGGTGATGATG 38
281 GGTCAGGTGATGATG 297
XX
RESULT 46
ABAS6889
ID ABAS6889 standard; DNA; 355 BP.
XX
AC ABAS6889;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #5194.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN MO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX
DR Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX Claim 1; SEQ ID NO 5194; 639bp + sequence listing; English.
PS
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;
XX
QY Query Match 13.7%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 22 GGTCAGGTGATGATG 38
281 GGTCAGGTGATGATG 297
XX
RESULT 47
ABA69454
ID ABA69454 standard; DNA; 355 BP.
XX
AC ABA69454;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #17759.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN MO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human foetal liver -
XX
XX Claim 4; SEQ ID NO 17759; 639bp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from [ftp.wipo.int/pub/published_pct_sequences](ftp:wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;

Query Match 13.7%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 GGTCGAGTGATGATG 38
DB 281 GGTCGAGTGATGATG 297

RESULT 48
ABA26508
ID ABA26508 standard; DNA; 355 BP.
XX
AC ABA26508;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #4974 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
PS Claim 1; SEQ ID No 4974; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC

CC at [ftp.wipo.int/pub/published_pct_sequences](ftp:wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;

Query Match 13.7%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 GGTCGAGTGATGATG 38
DB 281 GGTCGAGTGATGATG 297

RESULT 49
ABA36383
ID ABA36383 standard; DNA; 355 BP.
XX
AC ABA36383;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #14849 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
PS Claim 4; SEQ ID No 14849; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](ftp:wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;

Query Match 13.7%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 GGTCCAGTGATGATG 38
 |||||
 Db 281 GGTCCAGTGATGATG 297

RESULT 50

AAK04990
 ID AAK04990 standard; DNA; 355 BP.
 XX

AC AAK04990;
 XX

DT 05-NOV-2001 (first entry)
 XX

DE Human brain expressed single exon probe SEQ ID NO: 4981.
 XX

KM Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX

OS Homo sapiens.
 XX

PN WO200157275-A2.
 XX

PD 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001MO-US00667.
 XX

PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX

DR WPI; 2001-483446/52.
 XX

PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX

PS Example 4; SEQ ID NO: 4981; 650bp + Sequence Listing; English.
 XX

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system,
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 CC

SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;
 XX

Query Match 13.7%; Score 17; DB 22; Length 355;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 GGTCCAGTGATGATG 38
 |||||
 Db 281 GGTCCAGTGATGATG 297

Search completed: November 7, 2003, 11:58:14
 Job time : 127.47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 05:52:38 ; Search time 33.5936 Seconds
(without alignments)
1629.223 Million cell updates/sec

Title: US-09-509-779-3_COPY_141_264

Perfect score: 124
Sequence: 1 CGATACGTGCGCCATCTGCA.....ATTCTTCACACACTGCTGC 124

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	15.3	45	4	US-09-144-428-61
2	18	14.5	2405	1	US-08-454-097-30
3	18	14.5	2405	3	US-08-185-359-30
4	17	13.7	363	3	US-09-328-352-1220
5	17	13.7	3494	3	US-09-139-802-200
6	17	13.7	3494	4	US-09-659-786-200
7	17	13.7	8654	4	US-08-961-527-98
8	16	12.9	266	4	US-09-313-294A-3641
9	16	12.9	458	4	US-09-149-476-239
10	16	12.9	475	4	US-09-149-476-76
11	16	12.9	1059	3	US-08-163-919A-1
12	16	12.9	1059	5	PCT-US84-14073-1
13	16	12.9	5769	1	US-08-652-971-1
14	16	12.9	5769	2	US-08-991-258A-1
15	16	12.9	5769	3	US-08-769-399-1
16	16	12.9	5769	3	US-08-991-953A-1
17	15	12.1	196	3	US-08-663-082-1
18	15	12.1	275	4	US-09-313-294A-2588
19	15	12.1	819	1	US-08-309-182B-2
20	15	12.1	942	4	US-09-252-991A-11847
21	15	12.1	978	4	US-09-252-991A-11595
22	15	12.1	1137	4	US-09-252-991A-5529
23	15	12.1	1196	1	US-08-290-448A-71
24	15	12.1	1196	1	US-08-290-448A-71
25	15	12.1	1196	1	US-08-175-069A-71
26	15	12.1	1196	4	US-08-461-939B-71
27	15	12.1	1196	4	US-08-464-000-71

28	15	12.1	1292	4	US-09-205-258-189	Sequence 189, App
29	15	12.1	1314	3	US-09-025-059-2	Sequence 2, Appl
30	15	12.1	1368	1	US-08-290-448A-79	Sequence 79, Appl
31	15	12.1	1368	1	US-08-290-448A-79	Sequence 79, Appl
32	15	12.1	1368	1	US-08-175-069A-79	Sequence 79, Appl
33	15	12.1	1368	4	US-08-461-939B-79	Sequence 79, Appl
34	15	12.1	1368	4	US-08-464-000-79	Sequence 79, Appl
35	15	12.1	1426	3	US-09-121-425-3	Sequence 3, Appl
36	15	12.1	1426	4	US-09-634-493A-3	Sequence 3, Appl
37	15	12.1	1506	1	US-08-149-105-4	Sequence 4, Appl
38	15	12.1	1506	1	US-08-317-847-4	Sequence 4, Appl
39	15	12.1	1545	4	US-09-252-991A-5553	Sequence 5553, Ap
40	15	12.1	1605	4	US-09-252-991A-11780	Sequence 11780, A
41	15	12.1	1794	3	US-08-123-934A-9	Sequence 9, Appl
42	15	12.1	1794	5	PCT-US94-10080-9	Sequence 9, Appl
43	15	12.1	1914	4	US-09-252-991A-5670	Sequence 5670, Ap
44	15	12.1	2308	3	US-09-382-256-9	Sequence 9, Appl
45	15	12.1	2308	3	US-09-395-115-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-144-428-61/c
Sequence 61, Application US/09144428
Patent No. 6583108
GENERAL INFORMATION:
APPLICANT: BAYER CORPORATION, The
APPLICANT: TAMBURINI, Paul P
APPLICANT: DAVIS, Gary
APPLICANT: DELARIA, Katherine A
APPLICANT: MARJOR, Christopher W
APPLICANT: MULLER, Daniel K
TITLE OF INVENTION: HUMAN BIKUNIN
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive Suite 3200
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,428
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/03894
FILING DATE: 10-MAR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,106
FILING DATE: 11-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/019,793
FILING DATE: 14-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/725,251
FILING DATE: 04-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: CHAO, Mark
REGISTRATION NUMBER: 37,293
REFERENCE/DOCKET NUMBER: 96,223-II
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 913-0001
TELEFAX: (312) 913-0002
INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: not relevant
MOLECULE TYPE: DNA (genomic)
US-09-144-428-61

Query Match 15.3%; Score 19; DB 4; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AGTGGCCATCTGCAGG 23
|||||
DB 38 AGTGGCCATCTGCAGG 20

RESULT 2
US-08-454-097-30/C

Sequence 30, Application US/08454097
Patent No. 5686412
GENERAL INFORMATION:
APPLICANT: Hoechst, Merl F.
TITLE OF INVENTION: Protein Kinases
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,097
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/185,359
FILING DATE: 21-JAN-1994
APPLICATION NUMBER: US 08/008,001
FILING DATE: 21-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,783
FILING DATE: 03-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5686412and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31853
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2405 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 67..1197
US-08-454-097-30

Query Match 14.5%; Score 18; DB 1; Length 2405;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 GGGAGATGTATCATTC 108
|||||
DB 2241 GGGAGATGTATCATTC 2224

RESULT 3

US-08-185-359-30/C

Sequence 30, Application US/08185359
Patent No. 6060296
GENERAL INFORMATION:
APPLICANT: Hoechst, Merl F.
TITLE OF INVENTION: Protein Kinases
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,359
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,001
FILING DATE: 21-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,783
FILING DATE: 03-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 6060296and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31853
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2405 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 67..1197
US-08-185-359-30

Query Match 14.5%; Score 18; DB 3; Length 2405;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 GGGAGATGTATCATTC 108
|||||
DB 2241 GGGAGATGTATCATTC 2224

RESULT 4
US-09-328-352-1220
Sequence 1220, Application US/09328352
Patent No. 6562358
GENERAL INFORMATION:
APPLICANT: Gary L. Bretton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: CTC99-03PA

;/ CURRENT APPLICATION NUMBER: US/09/328,352
;/ CURRENT FILING DATE: 1999-06-04
;/ NUMBER OF SEQ ID NOS: 8252
;/ SEQ ID NO 1220
;/ LENGTH: 363
;/ TYPE: DNA
;/ ORGANISM: Acinetobacter baumannii
US-09-328-1220

Query Match 13.7%; Score 17; DB 4; Length 363;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TGTGTGTGTGTGTGGG 93
DB 213 TGTGTGTGTGTGTGGG 229

RESULT 5
US-09-139-802-200/c
;/ Sequence 200, Application US/09139802
;/ Patent No. 6180084
;/ GENERAL INFORMATION:
;/ APPLICANT: Ruoslahti, Erkki
;/ APPLICANT: Pasqualini, Renata
;/ TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
;/ TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
;/ TITLE OF INVENTION: Same
;/ FILE REFERENCE: P-LU 3203
;/ CURRENT APPLICATION NUMBER: US/09/139,802
;/ CURRENT FILING DATE: 1998-08-25
;/ EARLIER APPLICATION NUMBER: 08/926,914
;/ EARLIER FILING DATE: 1997-09-10
;/ EARLIER APPLICATION NUMBER: 08/710,067
;/ EARLIER FILING DATE: 1996-09-10
;/ NUMBER OF SEQ ID NOS: 226
;/ SOFTWARE: PatentIn Ver. 2.0
;/ SEQ ID NO 200
;/ LENGTH: 3494
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: (121)..(3024)
US-09-139-802-200

Query Match 13.7%; Score 17; DB 3; Length 3494;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CCATCTGCAGGGTCCAG 28
DB 1762 CCATCTGCAGGGTCCAG 1746

RESULT 6
US-09-659-786-200/c
;/ Sequence 200, Application US/09659786
;/ Patent No. 6491894
;/ GENERAL INFORMATION:
;/ APPLICANT: Ruoslahti, Erkki
;/ APPLICANT: Pasqualini, Renata
;/ TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
;/ TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
;/ TITLE OF INVENTION: Same
;/ FILE REFERENCE: P-LU 3203
;/ CURRENT APPLICATION NUMBER: US/09/659,786
;/ CURRENT FILING DATE: 2000-09-11
;/ PRIOR APPLICATION NUMBER: 08/926,914
;/ PRIOR FILING DATE: 1997-09-10
;/ PRIOR APPLICATION NUMBER: 08/710,067
;/ PRIOR FILING DATE: 1996-09-10
;/ NUMBER OF SEQ ID NOS: 226

;/ SOFTWARE: PatentIn Ver. 2.0
;/ SEQ ID NO 200
;/ LENGTH: 3494
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: (121)..(3024)
US-09-659-786-200

Query Match 13.7%; Score 17; DB 4; Length 3494;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CCATCTGCAGGGTCCAG 28
DB 1762 CCATCTGCAGGGTCCAG 1746

RESULT 7
US-08-961-527-98/c
;/ Sequence 98, Application US/08961527
;/ Patent No. 6420135
;/ GENERAL INFORMATION:
;/ APPLICANT: Charles Kunsch
;/ TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
;/ NUMBER OF SEQUENCES: 391
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Human Genome Sciences, Inc.
;/ STREET: 9410 Key West Avenue
;/ CITY: Rockville
;/ STATE: Maryland
;/ COUNTRY: USA
;/ ZIP: 20850
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
;/ COMPUTER: HP Vectra 486/33
;/ OPERATING SYSTEM: MSDOS version 6.2
;/ SOFTWARE: ASCII Text
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/961,527
;/ CLASSIFICATION: 424
;/ FILING DATE:
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER:
;/ FILING DATE:
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Brookes, A. Anders
;/ REGISTRATION NUMBER: 36,373
;/ REFERENCE/DOCKET NUMBER: PB340P1
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (301) 309-8504
;/ TELEFAX: (301) 309-8512
;/ INFORMATION FOR SEQ ID NO: 98:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 8654 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: linear
US-08-961-527-98

Query Match 13.7%; Score 17; DB 4; Length 8654;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 GACTGTGTGTGTGTCTG 90
DB 1105 GACTGTGTGTGTGTCTG 1089

RESULT 8
US-09-313-294A-3641/c
;/ Sequence 3641, Application US/09313294A

```
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Ialagudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 3641
/ LENGTH: 266
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700612036H1
/ LOCATION: 128
/ OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-3641

Query Match      12.9%; Score 16; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      47 AGATGTCAGCTGAAA 62
Db      155 AGATGTCAGCTGAAA 140

RESULT 9
US-09-149-476-239/c
/ Sequence 239, Application US/09149476
/ Patent No. 6420526
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 186 Human Secreted proteins
/ FILE REFERENCE: P2002P1
/ CURRENT APPLICATION NUMBER: US/09/149,476
/ CURRENT FILING DATE: 1998-09-08
/ EARLIER APPLICATION NUMBER: PCT/US98/04493
/ EARLIER FILING DATE: 1998-03-06
/ EARLIER APPLICATION NUMBER: 60/040,162
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,333
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/038,621
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,626
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,334
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,336
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,163
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/047,600
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,615
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,597
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,502
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,633
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,583
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,617
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,618
```

```
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,503
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,592
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,581
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,584
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,500
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,587
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,492
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,598
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,613
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,582
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,596
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,612
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,632
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,601
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/043,580
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,568
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,314
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,569
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,311
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,671
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,674
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,669
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,312
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/048,974
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/056,886
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,877
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,889
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,893
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,630
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,878
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,662
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,872
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,882
/ EARLIER FILING DATE: 1997-08-22
```

EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,539
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,553
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 12.9%; Score 16; DB 4; Length 458;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 TCTGCAGGTCACGT 30
DB 238 TCTGCAGGTCACGT 223

RESULT 10
US-09-149-476-76/c
Sequence 76, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OR INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23

[illegible]

EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 12.9%; Score 16; DB 4; Length 475;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 TCTGCGAGGTCGAGT 30
DB 226 TCTGCGAGGTCGAGT 211

RESULT 11

US-08-163-919A-1
Sequence 1, Application US/08163919A

GENERAL INFORMATION:
APPLICANT: THOMAS, Wayne R., CHUA, Kaw-Yan, ROGERS, Bruce L., and
APPLICANT: KUO, Mei-chang
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A HOUSE DUST MITE
TITLE OF INVENTION: ALLERGEN, DER P III, AND USES THEREFOR
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/163,919A
FILING DATE: 08-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IPC-072 (IMI-041)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 63..848
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 150..848

US-08-163-919A-1

Query Match 12.9%; Score 16; DB 3; Length 1059;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 AGCTGAAACAAACAA 70
DB 958 AGCTGAAACAAACAA 973

RESULT 12

PCT-US94-14073-1
Sequence 1, Application PC/TUS9414073

GENERAL INFORMATION:
APPLICANT: THOMAS, Wayne R., CHUA, Kaw-Yan, ROGERS, Bruce L., and
APPLICANT: KUO, Mei-chang
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A HOUSE DUST MITE
TITLE OF INVENTION: ALLERGEN, DER P III, AND USES THEREFOR
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/14073
FILING DATE: 08-DEC-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IPC-072 (IMI-041)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 63..848
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 150..848
PCT-US94-14073-1

Query Match 12.9%; Score 16; DB 5; Length 1059;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 AGCTGAAACAAACAA 70
DB 958 AGCTGAAACAAACAA 973

RESULT 13

US-08-652-971-1/C
Sequence 1, Application US/08652971

GENERAL INFORMATION:
APPLICANT: Cheng, Jili
APPLICANT: Lasky, Laurence A.
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE

TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd.
CITY: South San Francisco
STATE: California
COUNTRY: United States
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,971
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dregler, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P1033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 952-3216
TELEFAX: (415) 952-9881
TELEX: 910 371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5769 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 379..4686
US-08-652-971-1

Query Match 12.9%; Score 16; DB 1; Length 5769;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 TTGTGCTCTGGGAGAG 96
DB 890 TTGTGCTCTGGGAGAG 875

RESULT 14
US-08-991-258A-1/c
Sequence 1, Application US/08991258A
Patent No. 592887
GENERAL INFORMATION:
APPLICANT: Cheng, Jill
APPLICANT: Lasky, Laurence A.
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLHRR, HOBBACH, TEST, ALBRITTON & HERBERT, LLP
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,258A
FILING DATE: 17-DEC-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/652,971
FILING DATE: 24-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dregler, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63478-3/WHD/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5769 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 379..4686
US-08-991-258A-1

Query Match 12.9%; Score 16; DB 2; Length 5769;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 TTGTGCTCTGGGAGAG 96
DB 890 TTGTGCTCTGGGAGAG 875

RESULT 15
US-08-769-399-1/c
Sequence 1, Application US/08769399
Patent No. 5976852
GENERAL INFORMATION:
APPLICANT: Cheng, Jill
APPLICANT: Lasky, Laurence A.
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd.
CITY: South San Francisco
STATE: California
COUNTRY: United States
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,399
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dregler, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P1033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 225-3216
TELEFAX: (415) 952-9881
TELEX: 910 371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5769 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 379..4686
US-08-769-399-1

Query Match 12.9%; Score 16; DB 2; Length 5769;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 TTGTGCTGGGAG 96
|||||
Db 890 TTGTGCTGGGAG 875

RESULT 16
US-08-991-953A-1/C
Sequence 1, Application US/08991953A
Patent No. 6083748

GENERAL INFORMATION:
APPLICANT: Cheng, Jill
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
TITLE OF INVENTION: PHOSPHATASE, FTP LAMBDA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT, LLP
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,953A
FILING DATE: 16-DEC-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/652,971
FILING DATE: 24-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63478-3/WHD/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249

TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5769 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 379..4686
US-08-991-953A-1

Query Match 12.9%; Score 16; DB 3; Length 5769;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 TTGTGCTGGGAG 96
|||||
Db 890 TTGTGCTGGGAG 875

RESULT 17

US-08-663-082-1
Sequence 1, Application US/08663082
Patent No. 6043411

GENERAL INFORMATION:
APPLICANT: NISHIZAWA, Osamu
TITLE OF INVENTION: GENE FOR FATTY ACID DESATURASE, VECTOR
TITLE OF INVENTION: CONTAINING SAID GENE, PLANT TRANSFORMED WITH SAID GENE,
TITLE OF INVENTION: AND PROCESS FOR CREATING SAID PLANT
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,082
FILING DATE: 25-JUN-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/02288
FILING DATE: 28-DEC-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 93/352858
FILING DATE: 28-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 81356/107
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399

TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 196 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Anabaena variabilis
STRAIN: IAM M-3
US-08-663-082-1

Query Match 12.1%; Score 15; DB 3; Length 196;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 TTGTGCTGGGAG 95
|||||
Db 41 TTGTGCTGGGAG 55

RESULT 18
US-09-313-294A-2588/C
Sequence 2588, Application US/09313294A
Patent No. 6476212

GENERAL INFORMATION:
APPLICANT: Lalugudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600

Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 5529
LENGTH: 1137
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: unsure
LOCATION: (1131)
OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-5529

Query Match 12.1%; Score 15; DB 4; Length 1137;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TGGCCATCTGCAGG 22
DB 813 TGGCCATCTGCAGG 799

RESULT 23

US-08-290-448A-71
Sequence 71, Application US/08290448A
Patent No. 5676954

GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,448A
FILING DATE: August 15, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 1196 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1161
US-08-290-448A-71

Query Match 12.1%; Score 15; DB 1; Length 1196;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 AGAATGTATCATTC 108
DB 449 AGAATGTATCATTC 463

RESULT 24
US-08-290-448A-71
Sequence 71, Application US/08290448A
Patent No. 5698204

GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,448A
FILING DATE: August 15, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 1196 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1161
US-08-290-448A-71

Query Match 12.1%; Score 15; DB 1; Length 1196;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 AGAATGTATCATTC 108
DB 449 AGAATGTATCATTC 463

RESULT 25
US-08-175-069A-71
Sequence 71, Application US/08175069A
Patent No. 5776761
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,069A
FILING DATE: December 29, 1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IM1-018DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 1196 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1161
US-08-175-069A-71

Query Match 12.1%; Score 15; DB 1; Length 1196;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 94 AGAATGTAATCATTC 108
|||||
DB 449 AGAATGTAATCATTC 463

RESULT 26
US-08-461-939B-71
Sequence 71, Application US/08461939B
Patent No. 6335019
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Methods For Treating Sensitivity To A
NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,939B
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/464,000
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: US 08/290,448
FILING DATE: 15-AUG-1994
APPLICATION NUMBER: US 07/529,951
FILING DATE: 29-MAY-1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: 17-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IM1-018CNDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 1196 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1161
US-08-461-939B-71

Query Match 12.1%; Score 15; DB 4; Length 1196;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 94 AGAATGTAATCATTC 108
|||||
DB 449 AGAATGTAATCATTC 463

RESULT 27
US-08-464-000-71
Sequence 71, Application US/08464000
Patent No. 6335020
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Peptides from Ragweed Pollen
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,000
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/290,448
FILING DATE: 15-AUG-1994
APPLICATION NUMBER: US 07/529,951
FILING DATE: 29-MAY-1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: 17-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMT-018CN2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 1196 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1161
US-08-464-000-71

Query Match 12.1% Score 15; DB 4; Length 1196;
Best Local Similarity 100.0%; Pred.No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 AGAATGTAATCATTC 108
|||||
Db 449 AGAATGTAATCATTC 463

RESULT 28
US-09-205-258-189
Sequence 189, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 189
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
US-09-205-258-189

Query Match 12.1% Score 15; DB 4; Length 1292;
Best Local Similarity 100.0%; Pred.No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 CTGAACAACAACAG 71
|||||
Db 60 CTGAACAACAACAG 74

RESULT 29
US-09-025-059-2
; Sequence 2, Application US/09025059
; Patent No. 6075136
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Guejler, Karl J.
; TITLE OF INVENTION: PROSTATE-ASSOCIATED SERINE PROTEASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,059
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0481 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1314 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGUT110
; CLONE: 2723646
; US-09-025-059-2

Query Match 12.1%; Score 15; DB 3; Length 1314;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 CTGAACAACAAACAG 71
|||||
Db 77 CTGAACAACAAACAG 91

RESULT 30
US-08-290-448A-79
; Sequence 79, Application US/08290448A
; Patent No. 5676954
; GENERAL INFORMATION:
; APPLICANT: Rogers, Bruce
; APPLICANT: Klapper, David G.
; APPLICANT: Ratnar, Thorunn
; APPLICANT: Kuo, Mei-chang
; TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston

STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,448A
FILING DATE: August 15, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 1368 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1366
US-08-290-448A-79

Query Match 12.1%; Score 15; DB 1; Length 1368;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 AGAATGTATCATTC 108
|||||
Db 452 AGAATGTATCATTC 466

RESULT 31
US-08-290-448A-79
; Sequence 79, Application US/08290448A
; Patent No. 5698204
; GENERAL INFORMATION:
; APPLICANT: Rogers, Bruce
; APPLICANT: Klapper, David G.
; APPLICANT: Ratnar, Thorunn
; APPLICANT: Kuo, Mei-chang
; TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,448A
; FILING DATE: August 15, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/529,951
; FILING DATE: May 29, 1990

APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ. ID NO.: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 1368 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1366
US-08-290-448A-79

Query Match 12.1%; Score 15; DB 1; Length 1368;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 AGAATGTAATCATTC 108
DB 452 AGAATGTAATCATTC 466

RESULT 32
US-08-175-069A-79
Sequence 79, Application US/08175069A
Patent No. 5776761
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,069A
FILING DATE: December 29, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018DV
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ. ID NO.: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 1368 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1366
US-08-175-069A-79

Query Match 12.1%; Score 15; DB 1; Length 1368;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 AGAATGTAATCATTC 108
DB 452 AGAATGTAATCATTC 466

RESULT 33
US-08-461-939B-79
Sequence 79, Application US/08461939B
Patent No. 6335019
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Methods For Treating Sensitivity To A
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,939B
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/464,000
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: US 08/290,448
FILING DATE: 15-AUG-1994
APPLICATION NUMBER: US 07/529,951
FILING DATE: 29-MAY-1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: 17-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CNDV
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ. ID NO.: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 1368 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1366
US-08-461-939B-79

Query Match 12.1%; Score 15; DB 4; Length 1368;
Best Local Similarity 100.0%; Pred. No. 71;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 AGAATGTAATCATTC 108
|||||
Db 452 AGAATGTAATCATTC 466

RESULT 34
US-08-464-000-79

; Sequence 79, Application US/08464000

; Patent No. 6335020

; GENERAL INFORMATION:

; APPLICANT: Rogers, Bruce

; APPLICANT: Klapper, David G.

; APPLICANT: Rafnar, Thorunn

; APPLICANT: Kuo, Mei-chang

; TITLE OF INVENTION: Allergenic Peptides from Ragweed Pollen

; NUMBER OF SEQUENCES: 93

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 60 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/464,000

; FILING DATE: 05-JUN-1995

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/290,448

; FILING DATE: 15-AUG-1994

; APPLICATION NUMBER: US 07/529,951

; FILING DATE: 29-MAY-1990

; APPLICATION NUMBER: US 07/325,365

; FILING DATE: 17-MAR-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Amy E. Mandragouras

; REGISTRATION NUMBER: 36,207

; REFERENCE/DOCKET NUMBER: IMT-018CN2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)227-7400

; TELEFAX: (617)227-5941

; INFORMATION FOR SEQ ID NO: 79:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1368 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1366

; US-08-464-000-79

Query Match 12.1%; Score 15; DB 4; Length 1368;

Best Local Similarity 100.0%; Pred. No. 71;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 AGAATGTAATCATTC 108
|||||
Db 452 AGAATGTAATCATTC 466

RESULT 35
US-09-121-425-3

; Sequence 3, Application US/09121425

; Patent No. 6153418

; GENERAL INFORMATION:

; APPLICANT: Lehmann, Martin

; TITLE OF INVENTION: Consensus Phytases

; FILE REFERENCE: consensus phytases 13239

; CURRENT APPLICATION NUMBER: US/09/121,425

; CURRENT FILING DATE: 1998-07-23

; EARLIER APPLICATION NUMBER: EPO 97112688.3

; EARLIER FILING DATE: 1997-07-24

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1426

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:consensus

; US-09-121-425-3

Query Match 12.1%; Score 15; DB 3; Length 1426;

Best Local Similarity 100.0%; Pred. No. 70;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 ATGTCAAGCTGAAAA 63
|||||
Db 1250 ATGTCAAGCTGAAAA 1264

RESULT 36
US-09-634-493A-3

; Sequence 3, Application US/09634493A

; Patent No. 6579975

; GENERAL INFORMATION:

; APPLICANT: Lehmann, Martin

; TITLE OF INVENTION: Consensus Phytases

; FILE REFERENCE: consensus phytases 13239

; CURRENT APPLICATION NUMBER: US/09/634,493A

; CURRENT FILING DATE: 2000-08-08

; PRIOR APPLICATION NUMBER: US/09/121,425

; PRIOR FILING DATE: 1998-07-23

; PRIOR APPLICATION NUMBER: EPO 97112688.3

; PRIOR FILING DATE: 1997-07-24

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1426

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:consensus

; US-09-634-493A-3

Query Match 12.1%; Score 15; DB 4; Length 1426;

Best Local Similarity 100.0%; Pred. No. 70;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 ATGTCAAGCTGAAAA 63
|||||
Db 1250 ATGTCAAGCTGAAAA 1264

RESULT 37
US-08-149-105-4/C

; Sequence 4, Application US/08149105

; Patent No. 5538892

; GENERAL INFORMATION:

; APPLICANT: Donahoe, Patricia K.

; APPLICANT: Gustafson, Michael

; APPLICANT: He, Wei W.

; APPLICANT: Wang, Xiao-Fan

; TITLE OF INVENTION: TGF- TYPE I RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,105
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,673
; FILING DATE: March 11, 1993
; APPLICATION NUMBER: 07/853,396
; FILING DATE: March 18, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/211001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1506
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-149-105-4
;
Query Match 12.1%; Score 15; DB 1; Length 1506;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 ATCATTCTTCCACA 116
DB 902 ATCATTCTTCCACA 888

RESULT 38
US-08-317-847-4/c
; Sequence 4, Application US/08317847
; Patent No. 5547854
; GENERAL INFORMATION:
; APPLICANT: Donahoe, Patricia K.
; APPLICANT: Gustafson, Michael
; TITLE OF INVENTION: FOUR NOVEL RECEPTORS OF THE TGF-B
; TITLE OF INVENTION: FAMILY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,847
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: 08/029,673
; FILING DATE: March 11, 1993
; APPLICATION NUMBER: 07/853,396
; FILING DATE: March 18, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/127002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1506
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-317-847-4
;
Query Match 12.1%; Score 15; DB 1; Length 1506;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 ATCATTCTTCCACA 116
DB 902 ATCATTCTTCCACA 888

RESULT 39
US-09-252-991A-5553/c
; Sequence 5553, Application US/09252991A
; Patent No. 6551755
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27.
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5553
; TYPE: DNA
; LENGTH: 1545
; ORGANISM: Pseudomonas aeruginosa
;
US-09-252-991A-5553
;
Query Match 12.1%; Score 15; DB 4; Length 1545;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TGGCCATCTGCAGG 22
DB 1280 TGGCCATCTGCAGG 1266

RESULT 40
US-09-252-991A-11780
; Sequence 11780, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
```


PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11780
LENGTH: 1605
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11780

Query Match 12.1%; Score 15; DB 4; Length 1605;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 GGTGATGATGCTG 42
|||||
Db 409 GGTGATGATGCTG 423

RESULT 41

US-08-123-934A-9/C
Sequence 9, Application US/08123934A
Patent No. 6291206
GENERAL INFORMATION:
APPLICANT: MOZNEY, John
APPLICANT: CELESTE, Anthony J.
APPLICANT: THIES, R. Scott
APPLICANT: YAMAUTI, No. 62912060ru
TITLE OF INVENTION: RECEPTOR PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute Inc. - Legal Affairs
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,934A
FILING DATE: 17-SEP-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: LAZAR, Steven R
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5203
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 876 1170
TELEFAX: 617 876 5851
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1794 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: W-120
FEATURE:
NAME/KEY: CDS
LOCATION: 83..1591
US-08-123-934A-9

Query Match 12.1%; Score 15; DB 3; Length 1794;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 ATCATCTCTTCACA 116
|||||
Db 990 ATCATCTCTTCACA 976

RESULT 42

PCT-US94-10080-9/C
Sequence 9, Application PC/TUS9410080
GENERAL INFORMATION:
APPLICANT: GENETICS INSTITUTE, INC.
TITLE OF INVENTION: RECEPTOR PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute Inc. - Legal Affairs
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10080
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,934
FILING DATE: 17-SEP-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: LAZAR, Steven R
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5203-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-5851
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1794 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: W-120
FEATURE:
NAME/KEY: CDS
LOCATION: 83..1591
PCT-US94-10080-9

Query Match 12.1%; Score 15; DB 5; Length 1794;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 ATCATCTCTTCACA 116
|||||
Db 990 ATCATCTCTTCACA 976

RESULT 43

US-09-252-991A-5670
Sequence 5670, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
RILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 5670
LENGTH: 1914
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5670

Query Match
Best Local Similarity 100.0%; Pred.No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TGGCCATCTGCAGG 22
DB 209 TGGCCATCTGCAGG 223

RESULT 44
US-09-382-256-9/c
Sequence 9, Application US/09382256A
Patent No. 6207814
GENERAL INFORMATION:
APPLICANT: MIYAZONO, Kohei
FRANZEN, Petra
YAMASHITA, Hideoshi
HELDIN, Carl-Henrik
TITLE OF INVENTION: ACTIVIN RECEPTOR LIKE KINASES, PROTEINS
HAVING SERINE THREONINE KINASE DOMAINS,
AND THEIR USE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSER: Fulbright & Jaworski L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/382,256A
FILING DATE: 24-Aug-1999
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/02367
FILING DATE: No. 6207814ember 17, 1993
APPLICATION NUMBER: GB 9224057.1
FILING DATE: No. 6207814ember 17, 1992
APPLICATION NUMBER: GB 9304677.9
FILING DATE: March 8, 1993
APPLICATION NUMBER: GB 9304680.3
FILING DATE: March 8, 1993
APPLICATION NUMBER: 9311047.6
FILING DATE: May 28, 1993
APPLICATION NUMBER: 9313763.6
FILING DATE: July 2, 1993
APPLICATION NUMBER: 9316099.2
FILING DATE: August 3, 1993
APPLICATION NUMBER: 321344.5
FILING DATE: October 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 6207814man D. Hanson
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5298.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:

LENGTH: 2308 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Mouse
FEATURE:
NAME/KEY: CDS
LOCATION: 77..1585
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-382-256-9

Query Match
Best Local Similarity 100.0%; Pred.No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 ATCATTCCTTCACA 116
DB 984 ATCATTCCTTCACA 970

RESULT 45
US-09-395-115-9/c
Sequence 9, Application US/09395115
Patent No. 6271365
GENERAL INFORMATION:
APPLICANT: Miyazono, Kohei; DiJke, Peter Ten;
FRANZEN, Petra; Yamashita, Hideoshi; Heldin, Carl-Henrik
TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins And Their Use
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSER: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 Kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/395,115
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/436,265
FILING DATE: 30-October-1995
APPLICATION NUMBER: PCT/GB93/02367
FILING DATE: 17-No. 6271365ember-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9224057.1
FILING DATE: 17-No. 6271365ember-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9304677.9
FILING DATE: 8-March-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9304680.3
FILING DATE: 8-March-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9311047.6
FILING DATE: 28-May-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9313763.6
FILING DATE: 2-July-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9136099.2
FILING DATE: 3-August-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9321344.5
FILING DATE: 15-October-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kohlei, Vineet
REGISTRATION NUMBER: 37,003
REFERENCE/DOCKET NUMBER: LUD 5298
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2308 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Mouse
FEATURE:
NAME/KEY: CDS
LOCATION: 77..1585
US-09-395-115-9

Query Match 12.1%; Score 15; DB 3; Length 2308;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 102 ATCATTCTTCACCA 116
|||
Db 984 ATCATTCTTCACCA 970

RESULT 46
US-08-436-265-9/c
Sequence 9, Application US/08436265
Patent No. 6316217
GENERAL INFORMATION:
APPLICANT: Miyazono, Kohlei, DiJke, Peter Ten;
APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik
TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins
TITLE OF INVENTION: Having Serine Threonine Kinase Domains And Their Use
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,265
FILING DATE: 30-October-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/02367
FILING DATE: 17-No. 6316217ember-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9224057.1
FILING DATE: 17-No. 6316217ember-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9304677.9
FILING DATE: 8-March-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9304680.3
FILING DATE: 8-March-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9311047.6
FILING DATE: 28-May-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9313763.6
FILING DATE: 2-July-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9136099.2
FILING DATE: 3-August-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9321344.5
FILING DATE: 15-October-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kohlei, Vineet
REGISTRATION NUMBER: 37,003
REFERENCE/DOCKET NUMBER: LUD 5298
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2308 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Mouse
FEATURE:
NAME/KEY: CDS
LOCATION: 77..1585
US-08-436-265-9

Query Match 12.1%; Score 15; DB 4; Length 2308;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 102 ATCATTCTTCACCA 116
|||
Db 984 ATCATTCTTCACCA 970

RESULT 47
US-09-679-187-9/c
Sequence 9, Application US/09679187
Patent No. 6331621
GENERAL INFORMATION:
APPLICANT: Miyazono, Kohlei, DiJke, Peter Ten;
APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik
TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins
TITLE OF INVENTION: Having Serine Threonine Kinase Domains And Their Use
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/679,187
FILING DATE: 03-OCT-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/436,265
FILING DATE: 30-October-1995
PRIOR APPLICATION DATA:

```
APPLICATION NUMBER: PCT/GB93/02367
FILING DATE: 17-No. 633161ember-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9224057.1
FILING DATE: 17-No. 633161ember-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9304677.9
FILING DATE: 8-March-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9304680.3
FILING DATE: 8-March-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9311047.6
FILING DATE: 28-May-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9313763.6
FILING DATE: 2-July-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9136099.2
FILING DATE: 3-August-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9321344.5
FILING DATE: 15-October-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kohlel, Vineet
REGISTRATION NUMBER: 37,003
REFERENCE/DOCKET NUMBER: LUD 5298
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2308 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Mouse
FEATURE:
NAME/KEY: CDS
LOCATION: 77..1585
US-09-679-187-9

Query Match          12.1%; Score 15; DB 4; Length 2308;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 102 ATCATTCTTCCACA 116
DB 984 ATCATTCTTCCACA 970
```

```
RESULT 48
US-09-489-847-67
Sequence 67, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489,847
EARLIER FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
```

```
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: Patentn Ver. 2.0
SEQ ID NO 67
LENGTH: 2434
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (12)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (27)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (73)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (75)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (103)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (130)
OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-67
```

```
Query Match          12.1%; Score 15; DB 4; Length 2434;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 44 CTTAGATGTCAGCT 58
DB 31 CTTAGATGTCAGCT 45
```

```
RESULT 49
US-07-959-943-10/c
Sequence 10, Application US/07959943
Patent No. 5418162
GENERAL INFORMATION:
APPLICANT: Blakely, Randy D.
APPLICANT: Fremeau Jr., Robert T.
TITLE OF INVENTION: Serotonin Transporter cDNA
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Kenneth D. Sibley, Bell, Seltzer, Park
ADDRESSER: and
ADDRESSER: Gibson
STREET: Post Office Drawer 31107
CITY: Raleigh
STATE: No. 5418162th Carolina
COUNTRY: U.S.A.
ZIP: 27622
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: 943
APPLICATION NUMBER: US/07/959,943
FILING DATE: 19921014
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-881-3140
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2508 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
POSITION IN GENOME:
UNITS: 2278 basepairs
FEATURE:
NAME/KEY: CDS
LOCATION: 73..1962
US-07-959-943-10

Query Match 12.1%; Score 15; DB 1; Length 2508;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGGGTCACGTGATG 34
Db 737 AGGGTCACGTGATG 723

RESULT 50
US-08-551-459-3/c
Sequence 3, Application US/08551459
Patent No. 5821350
GENERAL INFORMATION:
APPLICANT: Huang, Yue
APPLICANT: Karatzas, Costas N.
APPLICANT: Lazaris-Karatzas, Anthoula
APPLICANT: Delaquis, Annick
TITLE OF INVENTION: ASPERGILLUS NIGER BETA GALACTOSIDASE
TITLE OF INVENTION: GENE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/551,459
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06632/005001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 3057 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
ORIGINAL SOURCE: Aspergillus niger
ORGANISM: Aspergillus niger
US-08-551-459-3

Query Match 12.1%; Score 15; DB 1; Length 3057;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 AGATGTCACCTGAA 61
Db 248 AGATGTCACCTGAA 234

Search completed: November 7, 2003, 11:59:34
Job time: 34.5936 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 11:11:37 / Search time 842.956 Seconds

(without alignments)
6163.455 Million cell updates/sec

Title: US-09-509-779-1_COPY_154_280

Perfect score: 127

Sequence: 1 GTGCGATACCTGTGCCATCT.....ATTCTTCACACACTGTGTC 127

Scoring table: OLIGO_NUC

Gapop 60.0, Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: GenEmbl:*

```

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	127	100.0	1125	10	BC011127	BC011127 Mus muscu
2	127	100.0	1140	6	BD096966	BD096966 SAG:apopt
3	127	100.0	1140	10	AF092877	AF092877 Mus muscu
4	127	100.0	1140	10	AC123935	AC123935 Mus muscu
5	69	54.3	120515	10	AL929565	AL929565 Mouse DNA
6	49	38.6	244851	2	AC118615	AC118615 Mus muscu
7	38	29.9	612	11	BV071691	BV071691 S212P6546
8	34	26.8	112321	2	AC103666	AC103666 Mus muscu
9	33	26.0	187353	2	AC118192	AC118192 Mus muscu
10	32	25.2	224	6	BD077705	BD077705 5'EST of
11	32	25.2	342	6	AX282592	AX282592 Sequence
12	32	25.2	342	9	AF142060	AF142060 Homo sapi
13	32	25.2	342	6	BD096967	BD096967 Homo sapi
14	32	25.2	754	6	BD096981	BD096981 SAG:apopt
15	32	25.2	754	6	BD096985	BD096985 SAG:apopt
16	32	25.2	754	6	BD096986	BD096986 SAG:apopt
17	32	25.2	754	6	BD096987	BD096987 SAG:apopt
18	32	25.2	754	6	BD096988	BD096988 SAG:apopt
19	32	25.2	754	6	BD096989	BD096989 SAG:apopt
20	32	25.2	754	6	BD096990	BD096990 SAG:apopt
21	32	25.2	754	6	BD096991	BD096991 SAG:apopt
22	32	25.2	754	6	BD096992	BD096992 SAG:apopt
23	32	25.2	754	6	BD096994	BD096994 SAG:apopt
24	32	25.2	754	6	BD096995	BD096995 SAG:apopt
25	32	25.2	754	6	BD096996	BD096996 SAG:apopt
26	32	25.2	754	6	BD096997	BD096997 SAG:apopt
27	32	25.2	754	6	BD096998	BD096998 SAG:apopt
28	32	25.2	754	6	BD096999	BD096999 SAG:apopt
29	32	25.2	754	6	BD097000	BD097000 SAG:apopt
30	32	25.2	754	6	BD097001	BD097001 SAG:apopt
31	32	25.2	754	6	BD097002	BD097002 SAG:apopt
32	32	25.2	754	6	BD097003	BD097003 SAG:apopt
33	32	25.2	754	6	BD097004	BD097004 SAG:apopt
34	32	25.2	754	6	BD097005	BD097005 SAG:apopt
35	32	25.2	754	6	BD097006	BD097006 SAG:apopt
36	32	25.2	754	6	BD097007	BD097007 SAG:apopt
37	32	25.2	754	6	BD097008	BD097008 SAG:apopt
38	32	25.2	754	6	BD097009	BD097009 SAG:apopt
39	32	25.2	754	6	BD097010	BD097010 SAG:apopt
40	32	25.2	754	6	BD097011	BD097011 SAG:apopt
41	32	25.2	754	6	BD097012	BD097012 SAG:apopt
42	32	25.2	754	6	BD097013	BD097013 SAG:apopt
43	32	25.2	754	6	BD097014	BD097014 SAG:apopt
44	32	25.2	754	6	BD097015	BD097015 SAG:apopt
45	32	25.2	754	6	BD097016	BD097016 SAG:apopt

ALIGNMENTS

RESULT 1
BC011127
LOCUS
DEFINITION
IMAGE:41941077, complete cds.
ACCESSION
BC011127
VERSION
BC011127.1 GI:15029807
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
1 (bases 1 to 1125)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
Altschul S.F., Zeeberg B., Buewog K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Stetschenko L., Marinina K., Farmer A.A., Rubin G.M., Hong L., Stetler-O'M., Soares M.B., Bonaldi M.F., Casavant T.L., Schetter T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carinici P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.U., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Bakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska J., Smilun D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	22388287	2 (bases 1 to 1125)	Strausberg R.	Submitted (25-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amgdc@cm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Navavati, A.N., Gibbs, R.A.	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov Series: IRAX Plate: 24 Row X Column: 5 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, similarity but not identity to protein. Location/Qualifiers 1. .1125 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:19076 IMAGE:4194107" /tissue_type="liver, normal, 5 month old male mouse." /clone_type="NCI CGAP_L19" /lab_host="DH10B" /note="vector: pcMV-SPORT6" 1. .1125 /gene="Rnf7" /note="synonym: SAG" /db_xref="locusID:19823" /db_xref="MGI:1337096" 12. .353 /codon_start=1 /product="Rnf7 protein" /protein_id="AAH1127.1" /db_xref="GI:15029808" /db_xref="locusID:19823" /translation="MADYDEGEPCVYSSHGSGAGSGGGRHPSLKKNNAVAMWSMDVECHTALCRQVWDACLCRCAENKQEDCVVWAECHNSFNNCMSLWVKNNSCPCLC	

BASE COUNT	304 a	219 c	286 g	316 t
ORIGIN	OODWVCRIGK"			
Query Match	100.0%;	Score 127;	DB 10;	Length 1125;
Best Local Similarity	100.0%;	Pred. No. 5.1e-67;		
Matches 127;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 GTGGCATTCACCTGTCATCTGCAGGGTCAGGTGATGATGCGCTTCGATGTCAGC 60			
DB	149 GTGGCATTCCTGTGTCATCTGCAGGGTCAGGTGATGATGCGCTTCGATGTCAGC 208			
QY	61 TGAATAACAGCAAGAGAGACTGTGTTGTGTGCTGTGGGAGAGTGAACATTCCTTCACAA 120			
DB	209 TGAATAACAGCAAGAGAGACTGTGTTGTGTGCTGTGGGAGAGTGAACATTCCTTCACAA 268			
QY	121 CTGCTGC 127			
DB	269 CTGCTGC 275			
RESULT 2				
LOCUS	BD096966	1140 bp	DNA	linear
DEFINITION	SAG:apoptosis sensitivity gene.			
ACCESSION	BD096966			
VERSION	BD096966.1	GI:22642554		
KEYWORDS	JP 2001526063-A/1.			
SOURCE	unidentified			
ORGANISM	unidentified			
REFERENCE	1 (bases 1 to 1140)			
AUTHORS	Sun, Y.			
TITLE	SAG:apoptosis sensitivity gene			
JOURNAL	Patent: JP 2001526063-A 1 18-DEC-2001;			
COMMENT	WARNER LAMBERT CO			
	OS Unidentified			
	PN JP 2001526063-A/1			
	PD 18-DEC-2001			
	PF 15-DEC-1998 JP 2000525451			
	PR 19-DEC-1997 US 60/068179, 11-SEP-1998 US 60/099840 PI			
	YI SUN			
	PC C12N15/00, A61K31/711, A61K38/00, A61K48/00, A61P17/02, A61P35/00,			
	PC A61P9/06			
	PC A61P93/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC			
	PC C12N5/10, C12Q1/68,			
	PC G01N33/50, G01N33/68, C12N15/00, A61K37/02, C12N5/00 CC			
	Strandedness: Double;			
	CC Topology: Linear;			
	CC /note = 'Mouse SAG'			
	FH Key			
	FT CDS			
	FT misc_feature			
	FT Location/Qualifiers			
	1..1140			
	/organism="unidentified"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:32644"			
FEATURES				
source				
BASE COUNT	302 a	224 c	287 g	327 t
ORIGIN				
Query Match	100.0%;	Score 127;	DB 6;	Length 1140;
Best Local Similarity	100.0%;	Pred. No. 5.1e-67;		
Matches 127;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 GTGGATACCTGTGCCATCTGCAGGGTCAGAGTATGATGATCCCTTCGATGTCAGC 60			
DB	154 GTGGATACCTGTGCCATCTGCAGGGTCAGAGTATGATGATCCCTTCGATGTCAGC 213			
QY	61 TGAATAACAGCAAGAGAGACTGTGTTGTGTGCTGTGGGAGAGTGAACATTCCTTCACAA 120			
DB	214 TGAATAACAGCAAGAGAGACTGTGTTGTGTGCTGTGGGAGAGTGAACATTCCTTCACAA 273			

QY 121 CTGCTGC 127
|||||
Db 274 CTGCTGC 280

RESULT 3
AF092877 1140 bp mRNA linear ROD 24-JUL-2001
LOCUS Mus musculus zinc RING finger protein SAG mRNA, complete cds.
DEFINITION AF092877
ACCESSION AF092877.1 GI:4588031
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1140)
AUTHORS Duan,H., Wang,Y., Avitram,M., Swaroop,M., Loo,J.A., Bian,J., Tian,Y., Mueller,T., Bisgaier,C.L. and Sun,Y.
TITLE SAG, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents
JOURNAL Mol. Cell. Biol. 19 (4), 3145-3155 (1999)
MEDLINE 99182502
PUBMED 10082581
REFERENCE 2 (bases 1 to 1140)
AUTHORS Sun,Y.
TITLE Alterations of SAG mRNA in human cancer cell lines: requirement for the RING finger domain for apoptosis protection
JOURNAL Carcinogenesis 20 (10), 1899-1903 (1999)
MEDLINE 99435944
PUBMED 10506102
REFERENCE 3 (bases 1 to 1140)
AUTHORS Swaroop,M., Bian,J., Avitram,M., Duan,H., Bisgaier,C.L., Loo,J.A. and Sun,Y.
TITLE Expression, purification, and biochemical characterization of SAG, a RING finger redox-sensitive protein
JOURNAL Free Radical Biol. Med. 27, 193-202 (1999)
MEDLINE 10506102
PUBMED 10506102
REFERENCE 4 (bases 1 to 1140)
AUTHORS Swaroop,M., Wang,Y., Miller,P., Duan,H., Jackoe,T., Madore,S.J. and Sun,Y.
TITLE Yeast homolog of human SAG/ROCC/Rbx2/Hrt2 is essential for cell growth, but not for germination: chip profiling implicates its role in cell cycle regulation
JOURNAL Oncogene 19 (24), 2855-2866 (2000)
MEDLINE 20309864
PUBMED 10851089
REFERENCE 5 (bases 1 to 1140)
AUTHORS Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.F., Zhang,H. and Sun,Y.
TITLE Promotion of S-phase entry and cell growth under serum starvation by SAG/ROCC/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL Mol. Carcinog. 30 (1), 37-46 (2001)
MEDLINE 21152847
PUBMED 11255262
REFERENCE 6 (bases 1 to 1140)
AUTHORS Sun,Y.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-1998) Department of Molecular Biology, Parke-Davis, 2800 Plymouth Rd, Ann Arbor, MI 48105, USA
FEATURES
source 1..1140
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6xCBA"
/db_xref="taxon:10090"
/sex="female"
/feature_type="1ung"
/dev_stage="6-8 weeks"
17..358
/note="redox sensitive, metal binding; expression protects cells from apoptosis induced by redox compounds"

/codon_start=1
/product="zinc RING finger protein SAG"
/protein_id="AA025961.1"
/db_xref="GI:4588032"
/translation="MADVEDGESEPCVLSSHSGSAGSKGSDKMFSLKKNNAVAMNSWD
VECDTCAICRVQVMDACLRQCAENKQEDVVMVGSCNHSFHNCCSLVWKNNRCPLC
QDDWVVRIGK"

BASE COUNT 302 a 224 c 287 g 327 t
ORIGIN

Query Match 100.0%; Score 127; DB 10; Length 1140;
Best Local Similarity 100.0%; Pred. No. 5.1e-67;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGATACCTGTCATCTGCAGGTCAGGTGATGATGCTGCTTCGATGTCAGC 60
|||||
Db 154 GGGGATACCTGTCATCTGCAGGTCAGGTGATGATGCTGCTTCGATGTCAGC 213
|||||

QY 61 TGAACAAGCAAGAGAGACTGTGTGTGTGGGAGAGAGTAAACATTCTCCACAA 120
|||||
Db 214 TGAACAAGCAAGAGAGACTGTGTGTGTGGGAGAGAGTAAACATTCTCCACAA 273
|||||

QY 121 CTGCTGC 127
|||||
Db 274 CTGCTGC 280

RESULT 4
AC123935 219370 bp DNA linear ROD 02-OCT-2002
LOCUS Mus musculus chromosome 13 clone RP23-354J3, complete sequence.
DEFINITION AC123935
ACCESSION AC123935
VERSION AC123935.4 GI:23462972
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 219370)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 219370)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 219370)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 219370)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT On Oct 2, 2002 this sequence version replaced gi:22748547.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
Project Information -----
Center project name: M BA0354J03

FEATURES
source 1..219370
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="13"


```

BASE COUNT      61709 a 45568 c 45395 g 66698 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 84; DB 10; Length 219370;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCATACCTGTGCTGAGGGTCCAGGTGATGATGCTGCTTCGATGTCAGC 60
DB 79010 GTGCATACCTGTGCTGAGGGTCCAGGTGATGATGCTGCTTCGATGTCAGC 79069
QY 61 TGAACAAGCAGACGACTGTGT 84
DB 79070 TGAACAAGCAGACGACTGTGT 79093

RESULT 5
LOCUS AL929565/c 120515 bp DNA linear ROD 16-APR-2003
DEFINITION Mouse DNA sequence from clone RP23-454N16 on chromosome 4, complete
ACCESSION AL929565
VERSION AL929565
KEYWORDS HTG.
SOURCE HTG.
ORGANISM Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 120515)
Andrew, R.
Direct Submission
Submitted (16-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 16, 2003 this sequence version replaced gi:29825605.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Emi, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-454N16 is
constructed by the group of Dieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pPaces3.6.
Location/Qualifiers
1..120515
/organism="Mus musculus"
/mol_type="genomic DNA"

```

```

BASE COUNT      34369 a 25840 c 26441 g 33865 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 69; DB 10; Length 120515;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 CTTCGATGTCAGCTGAAACAGACAGAGGACGTGTGTGCTGGGAGAGTAAAC 106
DB 57827 CTTCGATGTCAGCTGAAACAGACAGAGGACGTGTGTGCTGGGAGAGTAAAC 57768
QY 107 CATTCCTTC 115
DB 57767 CATTCCTTC 57759

RESULT 6
LOCUS AC118615/c 244851 bp DNA linear HTG 16-MAR-2003
DEFINITION Mus musculus clone RP24-91A16, WORKING DRAFT SEQUENCE, 7 unordered
pieces.
ACCESSION AC118615
VERSION AC118615.4 GI:28975088
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE HTG.
ORGANISM Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 244851)
Birren, B., Nussbaum, C. and Lander, E.
Mus musculus, clone RP24-91A16
Unpublished
2 (bases 1 to 244851)
Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, A., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Grinde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
Minova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Punthang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tefaye, S., Theodore, J.,
Topham, K., Travars, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, D., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (19-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02144, USA
3 (bases 1 to 244851)
Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Archchidi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Deatellano, K.,
Diaz, J.S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,

```


Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

47 CTTGATGTCAGCTGAAACAGAGAGACTGTGT 84
|||||
407 CTTGATGTCAGCTGAAACAGAGAGACTGTGT 444
|||||

RESULT 8
AC103666 112321 bp DNA linear HTG 20-DEC-2002
LOCUS Mus musculus clone RP23-397H13, LOW-PASS SEQUENCE SAMPLING.
AC103666.2 GI:27311471
DEFINITION HTG, HTGS_PHASE0.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 112321)
Birren, B., Nusbaum, C., and Lander, E.
Mus musculus, clone RP23-397H13
Unpublished

2 (bases 1 to 112321)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barina, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cook, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McKernan, P., McKernan, K., McPheters, R., Meldrim, J.,
Menus, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. D., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (29-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 112321)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barina, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepe, Y., Collymore, A.,
Cook, A., Cook, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Graham, L., Grand-Pierre, N., Hafet, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Mihova, T.,
Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T.,
Mienna, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (20-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 20, 2002 this sequence version replaced gi:1719550.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)

http://fpc.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: U19615
Center clone name: 397_H_13

* NOTE: This record contains 89 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely for
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 1147: contig of 1147 bp in length
* 1148 1247: gap of 100 bp
* 1248 1248: contig of 1101 bp in length
* 2349 2448: gap of 100 bp
* 2449 3511: contig of 1063 bp in length
* 3512 3611: gap of 100 bp
* 3612 4746: contig of 1135 bp in length
* 4747 4846: gap of 100 bp
* 4847 6031: contig of 1184 bp in length
* 6031 6131: gap of 100 bp
* 6131 7290: contig of 1160 bp in length
* 7291 7390: gap of 100 bp
* 7391 8577: contig of 1187 bp in length
* 8578 8677: gap of 100 bp
* 8678 9937: contig of 1260 bp in length
* 9938 10037: gap of 100 bp
* 10038 11192: contig of 1155 bp in length
* 11193 11292: gap of 100 bp
* 11293 12454: contig of 1162 bp in length
* 12455 12554: gap of 100 bp
* 12555 13718: contig of 1164 bp in length
* 13719 13818: gap of 100 bp
* 13819 15002: contig of 1184 bp in length
* 15003 15102: gap of 100 bp
* 15103 16234: contig of 1132 bp in length
* 16235 16334: gap of 100 bp
* 16335 17461: contig of 1127 bp in length
* 17462 17561: gap of 100 bp
* 17562 18650: contig of 1089 bp in length
* 18651 18750: gap of 100 bp
* 18751 19912: contig of 1162 bp in length
* 19913 20012: gap of 100 bp
* 20013 21199: contig of 1187 bp in length
* 21200 21299: gap of 100 bp
* 21299 22462: contig of 1163 bp in length
* 22463 22562: gap of 100 bp
* 22563 23720: contig of 1158 bp in length
* 23721 23821: gap of 100 bp
* 23822 25005: contig of 1185 bp in length
* 25006 25105: gap of 100 bp
* 25106 26265: contig of 1160 bp in length
* 26266 26365: gap of 100 bp
* 26366 27528: contig of 1163 bp in length
* 27529 27628: gap of 100 bp
* 27629 28788: contig of 1160 bp in length
* 28789 28888: gap of 100 bp
* 28889 30034: contig of 1146 bp in length
* 30035 30134: gap of 100 bp
* 30135 31236: contig of 1102 bp in length
* 31237 31337: gap of 100 bp
* 31337 32513: contig of 1176 bp in length
* 32513 32612: gap of 100 bp

```

* 32613 33813: contig of 1201 bp in length
* 33914 33913: gap of 100 bp
* 35107 35107: contig of 1194 bp in length
* 35207 35207: gap of 100 bp
* 35208 35208: contig of 1224 bp in length
* 36432 36431: gap of 100 bp
* 36532 36531: contig of 1170 bp in length
* 37702 37701: gap of 100 bp
* 37802 37801: contig of 1166 bp in length
* 38967 38967: gap of 100 bp
* 39068 40235: contig of 1168 bp in length
* 40236 40335: gap of 100 bp
* 40336 41492: contig of 1157 bp in length
* 41493 41592: gap of 100 bp
* 41593 42691: contig of 1039 bp in length
* 42692 42791: gap of 100 bp
* 42792 43970: contig of 1179 bp in length
* 43971 44070: gap of 100 bp
* 44071 45150: contig of 1080 bp in length
* 45151 45250: gap of 100 bp
* 45251 46347: contig of 1037 bp in length
* 46348 46447: gap of 100 bp
* 46448 47608: contig of 1161 bp in length
* 47609 47708: gap of 100 bp
* 47709 48881: contig of 1173 bp in length
* 48882 48981: gap of 100 bp
* 50240 50240: contig of 1259 bp in length
* 50241 50340: gap of 100 bp
* 50341 51544: contig of 1204 bp in length
* 51545 51644: gap of 100 bp
* 51645 52828: contig of 1184 bp in length
* 52829 52928: gap of 100 bp
* 52929 54157: contig of 1229 bp in length
* 54158 54257: gap of 100 bp
* 54258 55450: contig of 1193 bp in length
* 55451 55550: gap of 100 bp
* 55551 56731: contig of 1181 bp in length
* 56732 56831: gap of 100 bp
* 56832 58019: contig of 1188 bp in length
* 58020 58119: gap of 100 bp
* 58120 59285: contig of 1166 bp in length
* 59286 59385: gap of 100 bp
* 59386 60525: contig of 1140 bp in length
* 60526 60625: gap of 100 bp
* 60626 61717: contig of 1032 bp in length
* 61718 61817: gap of 100 bp
* 61818 62991: contig of 1174 bp in length
* 62992 63091: gap of 100 bp
* 63092 64267: contig of 1176 bp in length
* 64268 64367: gap of 100 bp
* 64368 65552: contig of 1185 bp in length
* 65553 65652: gap of 100 bp
* 65653 66853: contig of 1201 bp in length
* 66854 66953: gap of 100 bp
* 66954 68145: contig of 1192 bp in length
* 68146 68245: gap of 100 bp
* 68246 69423: contig of 1178 bp in length
* 69424 69523: gap of 100 bp
* 69524 70737: contig of 1214 bp in length
* 70738 70837: gap of 100 bp
* 70838 71997: contig of 1160 bp in length
* 71998 72097: gap of 100 bp
* 72098 73254: contig of 1157 bp in length

```

Query Match 26.8%; Score 34; DB 2; Length 112321;
 Best Local Similarity 100.0%; Pred. No. 2.1e-09;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

90 TCTGGGAGAGTGTACATCTCTCCACACTG 123
 Db 78439 TCTGGGAGAGTGTACATCTCTCCACACTG 78472

RESULT 9
 AC118192/c
 LOCUS
 DEFINITION Mus musculus clone RP23-200C17, WORKING DRAFT SEQUENCE, 10 ordered
 pieces.
 ACCESSION AC118192
 VERSION AC118192.1 GI:20147957
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 187353)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Mus musculus, clone RP23-200C17.
 JOURNAL
 TITLE Unpublished
 REFERENCE 2 (bases 1 to 187353)
 AUTHORS
 REFERENCES
 AUTHORS

Unpublished
 2 (bases 1 to 187353)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gold, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kanat, A., Karatas, A., Kelle, C., LaRoque, K., Lamazares, R.,
 Landers, T., Lehotzky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., Macdonald, P., Major, U., Margulis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnappack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vasilev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, W.
 Direct Submission
 Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A. F. A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 Project Information
 Center project name: L19568
 Center clone name: 200_C_17
 Sequencing vector: Plasmid; N/A; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 184259 bases at least Q40
 Consensus quality: 185710 bases at least Q30
 Consensus quality: 186169 bases at least Q20
 Insert size: 194000; agarose-fp
 Insert size: 186453; sum-of-contigs
 Quality coverage: 7.4 in Q20 bases; agarose-fp
 Quality coverage: 7.7 in Q20 bases; sum-of-contigs
 NOTE: This is a 'working draft' sequence. It currently
 consists of 10 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces
 is believed to be correct as given, however the sizes
 of the gaps between them are based on estimates that have
 been provided by the submitter.
 This sequence will be replaced
 by the finished sequence as soon as it is available and

```
* the accession number will be preserved.
* 1 629: contig of 629 bp in length
* 630 729: gap of 100 bp
* 730 750: contig of 21 bp in length
* 751 850: gap of 100 bp
* 851 3212: contig of 2362 bp in length
* 3213 3312: gap of 100 bp
* 3313 6908: contig of 3596 bp in length
* 6909 7008: gap of 100 bp
* 7009 14434: contig of 7426 bp in length
* 14435 14534: gap of 100 bp
* 14535 32181: contig of 17647 bp in length
* 32182 32281: gap of 100 bp
* 32282 51167: contig of 18886 bp in length
* 51168 51267: gap of 100 bp
* 51268 72870: contig of 21603 bp in length
* 72871 72970: gap of 100 bp
* 72971 149992: contig of 77922 bp in length
* 149993 150092: gap of 100 bp
* 150093 187353: contig of 37261 bp in length.
FEATURES
source
1..187353
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-200C17"
/clone_1fb="RPC1-23 Female Mouse BAC"
1..629
/misc_feature
/note="assembly_fragment"
730..750
/note="assembly_fragment"
/misc_feature
/note="assembly_fragment"
851..3212
/note="assembly_fragment"
/misc_feature
/note="assembly_fragment"
3313..6908
/note="assembly_fragment"
7009..14434
/note="assembly_fragment"
/misc_feature
/note="assembly_fragment"
14535..32181
/note="assembly_fragment"
32282..51167
/note="assembly_fragment"
51268..72870
/note="assembly_fragment"
/misc_feature
/note="assembly_fragment"
72971..149992
/note="assembly_fragment"
150093..187353
/note="assembly_fragment"
clone_end:17
vector_side:right"
BASE COUNT 60628 a 34855 c 36129 g 54836 t 905 others
ORIGIN
Query Match 26.0%; Score 33; DB 2; Length 187353;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 13 TGCATCTGCAGGCTCCAGTGATGATGCTG 45
Db 46105 TGCATCTGCAGGCTCCAGTGATGATGCTG 46073
RESULT 10 224 bp DNA linear PAT 27-AUG-2002
LOCUS BD077705
DEFINITION 5'EST of secreted protein expressed in muscles and other mesodermal
tissues.
ACCESSION BD077705
VERSION BD077705.1 GI:22623308
KEYWORDS JP 2001512016-A/291.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
```

```
REFERENCE 1 (bases 1 to 224)
AUTHORS Edwards,J.B.D.M., Duclet,A. and Lacroix,B.
TITLE 5'EST of secreted protein expressed in muscles and other mesodermal
tissues
JOURNAL Parent: JP 2001512016-A 291 21-AUG-2001;
COMMENT
OS Homo sapiens (human)
PN JP 2001512016-A/291
PD 21-AUG-2001
PF 31-JUL-1998 JP 2000505295
PR 01-AUG-1997 US 08/905134
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLET, BRUNO PI
LACROIX
PC C12N15/09,C12N15/09,C07K14/47,C12M1/00,C12N15/00,C12N15/00 CC
blastn
CC identity 97
CC region 93..226
CC id W81645
CC est
CC blastn
CC identity 95
CC region 31..95
CC id W81645
CC est
CC blastn
CC identity 97
CC region 62..195
CC id W06951
CC est
CC blastn
CC identity 96
CC region 7..64
CC id W06951
CC est
CC blastn
CC identity 97
CC region 2..179
CC id W38711
CC est
CC Von Heljne matrix
CC score 3.7
CC seq ETCFLASHSSG/SK
CC n=a, g, c or t
FH Key Location/Qualifiers
FT misc_feature 89..222
FT misc_feature 26..90
FT misc_feature 89..222
FT misc_feature 33..90
FT misc_feature 45..222
FT sig_peptide 24..86
FT misc_feature 3.
FEATURES
source
1..224
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 41 a 60 c 83 g 38 t 2 others
ORIGIN
Query Match 25.2%; Score 32; DB 6; Length 224;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 14 GCCATCTGCAGGCTCCAGTGATGATGCTG 45
Db 174 GCCATCTGCAGGCTCCAGTGATGATGCTG 205
RESULT 11 342 bp DNA linear PAT 02-NOV-2001
LOCUS AX282592
DEFINITION Sequence 7 from Patent WO0175145.
ACCESSION AX282592
```

VERSION AX282592.1 GI:16609675
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Issakant,S.D., Huang,J., Sheung,J. and Pray,T.R.
TITLE Ubiquitin ligase assay
JOURNAL Patent: WO 0175145-A 7 11-OCT-2001;
Rigel Pharmaceuticals, Inc. (US)
FEATURES
source 1..342
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 82 a 84 c 108 g 68 t
ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 342;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GCCATCTGCAGGTCAGGTGATGATGCTG 45
151 GCCATCTGCAGGTCAGGTGATGATGCTG 182

Db 151 GCCATCTGCAGGTCAGGTGATGATGCTG 182

RESULT 12
AF142060 342 bp mRNA linear PRI 14-JUL-1999
LOCUS Homo sapiens RING finger protein (ROC2) mRNA, complete cds.
DEFINITION AF142060
ACCESSION AF142060
VERSION AF142060.1 GI:4809217
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 342)
AUTHORS Ohta,T., Michel,J., Schotelius,A.J. and Xiong,Y.
TITLE ROC2, a homolog of APC11, represents a family of cullin partners
JOURNAL Mol. Cell 3 (4), 535-541 (1999)
MEDLINE 99247022
PUBMED 10230407
REFERENCE 2 (bases 1 to 342)
AUTHORS Ohta,T., Michel,J. and Xiong,Y.
TITLE Direct Submission
JOURNAL Submitted (10-APR-1999) Linberger Comprehensive Cancer Center,
University of North Carolina at Chapel Hill, Mason Farm Rd. and
Manning Dr., Chapel Hill, NC 27599-7295, USA
FEATURES
source 1..342
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..342
/gene="ROC2"
1..342
/gene="ROC2"
/note="Cullin partnering protein with associated ubiquitin
ligase activity"
/ligase_activity
/codon_start=1
/product="RING finger protein"
/protein_id="AAD30147.1"
/db_xref="GI:4809218"
/translation="MADVDGEETCALASHSSGSGSGDKMFLSKKNAYAMWSMD
VECDICAIQRYQVMDACLRQCAENKQEDCVVWVGECNHSFNHCNLSWVKQNNRCPLC
QODWVQRIQK"
BASE COUNT 82 a 83 c 109 g 68 t
ORIGIN

Query Match 25.2%; Score 32; DB 9; Length 342;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GCCATCTGCAGGTCAGGTGATGATGCTG 45
151 GCCATCTGCAGGTCAGGTGATGATGCTG 182

Db 151 GCCATCTGCAGGTCAGGTGATGATGCTG 182

RESULT 13
BT007348 342 bp mRNA linear PRI 13-MAY-2003
LOCUS Homo sapiens ring finger protein 7 mRNA, complete cds.
DEFINITION BT007348
ACCESSION BT007348.1 GI:30583534
VERSION BT007348.1 GI:30583534
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 342)
AUTHORS Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,D., Lin,Y.,
Pheilan,M. and Farmer,A.
TITLE Cloning of human full-length CDSs in BD Creator(TM) system donor
vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 342)
AUTHORS Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,D., Lin,Y.,
Pheilan,M. and Farmer,A.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USA
COMMENT This CDS clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the
Harvard Institute of Proteomics. Each CDS has been cloned in two
forms: with and without stop-codon (to allow fusion with C-terminal
tag). The CDS has been directionally cloned using BD In-Fusion(TM)
cloning system between the SalI and HindIII sites of the pDNR-DUAL
vector. Additional sequences in the clone: 'ACC' after SalI site
and before 'ATG' to provide Kozak consensus sequence; 'GG' after
last codon and before HindIII site to maintain reading frame.
Clone distribution: <http://bioinfo.clontech.com/orfclones>.
FEATURES
source 1..342
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GH00912X1.0"
/clone_lib="BD Creator(TM) CDS Library derived from MGC
collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"
1..342
/codon_start=1
/product="ring finger protein 7"
/protein_id="AAP36012.1"
/db_xref="GI:30583535"
/translation="MADVDGEETCALASHSSGSGSGDKMFLSKKNAYAMWSMD
VECDICAIQRYQVMDACLRQCAENKQEDCVVWVGECNHSFNHCNLSWVKQNNRCPLC
QODWVQRIQK"
BASE COUNT 82 a 83 c 109 g 68 t
ORIGIN

Query Match 25.2%; Score 32; DB 9; Length 342;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GCCATCTGCAGGTCAGGTGATGATGCTG 45
151 GCCATCTGCAGGTCAGGTGATGATGCTG 182

Db 151 GCCATCTGCAGGTCAGGTGATGATGCTG 182

```

RESULT 14
BD096967 754 bp DNA linear PAT 27-AUG-2002
LOCUS BD096967 SAG:apoptosis sensitivity gene.
DEFINITION BD096967
ACCESSION BD096967.1 GI:22642555
VERSION JP 2001526063-A/2.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 2 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/2
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
A61P39/06,
C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC /note = 'Human SAG'
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..754.
FT misc_feature 1..754.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 205 a 155 c 201 g 193 t
ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45
151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182
DB

RESULT 15
BD096981 754 bp DNA linear PAT 27-AUG-2002
LOCUS BD096981 SAG:apoptosis sensitivity gene.
DEFINITION BD096981
ACCESSION BD096981.1 GI:22642569
VERSION JP 2001526063-A/16.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 16 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/16
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
A61P39/06,
C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC /note = 'Human SAG'
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..754.
FT misc_feature 1..754.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 205 a 155 c 201 g 193 t
ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45
151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182
DB

```

```

PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..754.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t
ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45
151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182
DB

RESULT 16
BD096985 754 bp DNA linear PAT 27-AUG-2002
LOCUS BD096985 SAG:apoptosis sensitivity gene.
DEFINITION BD096985
ACCESSION BD096985.1 GI:22642573
VERSION JP 2001526063-A/20.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 20 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/20
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
A61P39/06,
C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC /note = 'Human SAG'
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..754.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t
ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45
151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182
DB

```

Db 151 GCCATCTGCAGGTCACAGTGATGATGCCTG 182

RESULT 17
LOCUS BD096986 754 bp DNA linear PAT 27-AUG-2002
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096986.1 GI:22642574
VERSION JP 2001526063-A/21.
KEYWORDS JP 2001526063-A/21.
SOURCE unclassified
ORGANISM unclassified

REFERENCE
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 21 18-DEC-2001;
WARNER LAMBERT CO

COMMENT
OS Unidentified
PN JP 2001526063-A/21
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339
FT Location/Qualifiers

FEATURES
source 1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 207 a 154 c 201 g 192 t

ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 151 GCCATCTGCAGGTCACAGTGATGATGCCTG 182

RESULT 18
LOCUS BD096987 754 bp DNA linear PAT 27-AUG-2002
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096987.1 GI:22642575
VERSION JP 2001526063-A/22.
KEYWORDS JP 2001526063-A/22.
SOURCE unclassified
ORGANISM unclassified

REFERENCE
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 22 18-DEC-2001;
WARNER LAMBERT CO

COMMENT
OS Unidentified
PN JP 2001526063-A/22
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339
FT Location/Qualifiers

FEATURES
source 1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 153 c 202 g 193 t

ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 151 GCCATCTGCAGGTCACAGTGATGATGCCTG 182

RESULT 19
LOCUS BD096988 754 bp DNA linear PAT 27-AUG-2002
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096988.1 GI:22642576
VERSION JP 2001526063-A/23.
KEYWORDS JP 2001526063-A/23.
SOURCE unclassified
ORGANISM unclassified

REFERENCE
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 23 18-DEC-2001;
WARNER LAMBERT CO

COMMENT
OS Unidentified
PN JP 2001526063-A/23
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339
FT Location/Qualifiers

FEATURES
source 1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t

ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 GCCATCTGCAGGTCACAGTGATGATGCCTG 45

Db 151 GCCATCTGCAGGTCACAGGTGATGATGCTTG 182

RESULT 20
LOCUS BD096989 754 bp DNA linear PAT 27-AUG-2002
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096989
VERSION BD096989.1 GI:22642577
KEYWORDS JP 2001526063-A/24.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 24 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/24
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t

ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GCCATCTGCAGGTCACAGGTGATGATGCTTG 45
151 GCCATCTGCAGGTCACAGGTGATGATGCTTG 182

Db

RESULT 21
LOCUS BD096990 754 bp DNA linear PAT 27-AUG-2002
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096990
VERSION BD096990.1 GI:22642578
KEYWORDS JP 2001526063-A/25.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 25 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/25
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI

FEATURES
source
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t

ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GCCATCTGCAGGTCACAGGTGATGATGCTTG 45
151 GCCATCTGCAGGTCACAGGTGATGATGCTTG 182

Db

RESULT 22
LOCUS BD096991 754 bp DNA linear PAT 27-AUG-2002
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096991
VERSION BD096991.1 GI:22642579
KEYWORDS JP 2001526063-A/26.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 26 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/26
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t

ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 208 a 152 c 202 g 192 t

ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GCCATCTGCAGGTCACAGGTGATGATGCTTG 45
151 GCCATCTGCAGGTCACAGGTGATGATGCTTG 182

Db

RESULT 23
LOCUS BD096992 754 bp DNA linear PAT 27-AUG-2002
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096992
VERSION BD096992.1 GI:22642580
KEYWORDS JP 2001526063-A/27.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 27 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/27
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t

ORIGIN

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 45
|||||
Db 151 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 182

RESULT 23
BD096992 754 bp DNA linear PAT 27-AUG-2002
LOCUS SAG:apoptosis sensitivity gene.
DEFINITION BD096992
ACCESSION BD096992.1 GI:22642580
VERSION JP 2001526063-A/27.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 27 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/27
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 45
|||||
Db 151 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 182

RESULT 24
BD096994 754 bp DNA linear PAT 27-AUG-2002
LOCUS SAG:apoptosis sensitivity gene.
DEFINITION BD096994
ACCESSION BD096994.1 GI:22642582
VERSION JP 2001526063-A/29.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 29 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/29
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451

PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 207 a 155 c 201 g 191 t

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 45
|||||
Db 151 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 182

RESULT 25
BD096995 754 bp DNA linear PAT 27-AUG-2002
LOCUS SAG:apoptosis sensitivity gene.
DEFINITION BD096995
ACCESSION BD096995.1 GI:22642583
VERSION JP 2001526063-A/30.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 30 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/30
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FT Location/Qualifiers
1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t

Query Match 25.2%; Score 32; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGTCGAGTGATGATGCTG 45
 DB 151 GCCATCTGCAGGTCGAGTGATGATGCTG 182

RESULT 26
 AF092878
 LOCUS AF092878 754 bp mRNA linear PRI 24-JUL-2001
 DEFINITION Homo sapiens zinc RING finger protein SAG mRNA, complete cds.
 ACCESSION AF092878
 VERSION AF092878.1 GI:4588033
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
 1 (bases 1 to 754)
 Tian, Y., Wang, Y., Aviram, M., Swaroop, M., Loo, J.A., Bian, J., SAG, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents
 Mol. Cell. Biol. 19 (4), 3145-3155 (1999)
 99182502
 10082581
 2 (bases 1 to 754)
 Sun, Y.
 Alterations of SAG mRNA in human cancer cell lines: requirement for the RING finger domain for apoptosis protection
 Carcinogenesis 20 (10), 1899-1903 (1999)
 99435944
 10506102
 3 (bases 1 to 754)
 Swaroop, M., Bian, J., Aviram, M., Duan, H., Bisgaier, C.L., Loo, J.A. and Sun, Y.
 Expression, purification, and biochemical characterization of SAG, a RING finger redox-sensitive protein
 Free Radical Biol. Med. 27, 193-202 (1999)
 4 (bases 1 to 754)
 Swaroop, M., Wang, Y., Miller, P., Duan, H., Jackoe, T., Madore, S.J. and Sun, Y.
 Yeast homolog of human SAG/ROC2/Rbx2/Hrt2 is essential for cell growth, but not for germination: chip profiling implicates its role in cell cycle regulation
 Oncogene 19 (24), 2855-2866 (2000)
 20309864
 10851089
 5 (bases 1 to 754)
 Duan, H., Tsvetkov, L.M., Liu, Y., Song, Y., Swaroop, M., Wen, R., Kung, H.F., Zhang, H. and Sun, Y.
 Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: a association with inhibition of p27 accumulation
 Mol. Carcinog. 30 (1), 37-46 (2001)
 21152847
 11255262
 6 (bases 1 to 754)
 Sun, Y.
 Direct Submision
 Submitted (16-SEP-1998) Department of Molecular Biology, Parke-Davis, 2800 Plymouth Rd, Ann Arbor, MI 48105, USA

FEATURES
 SOURCE
 1..754
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3q22-q24"
 /cell_line="Hela, D98/AH-2, HPRT-"
 1..342
 /function="growth promotion"
 /note="redox sensitive, metal binding; expression protects cells from apoptosis induced by redox compounds"

CDS

BASE COUNT 205 a 155 c 201 g 193 t

ORIGIN

Query Match
 Best Local Similarity 100.0%; Pred. No. 4e-08;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGTCGAGTGATGATGCTG 45
 DB 151 GCCATCTGCAGGTCGAGTGATGATGCTG 182

RESULT 27
 BC008627
 LOCUS BC008627 816 bp mRNA linear PRI 12-JUL-2001
 DEFINITION Homo sapiens, ring finger protein 7, clone MGC:17274 IMAGE:4177613, mRNA, complete cds.
 ACCESSION BC008627
 VERSION BC008627.1 GI:14250388
 KEYWORDS
 SOURCE MGC.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
 1 (bases 1 to 816)
 Strausberg, R.
 Direct Submission
 Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: villalona@bcm.tmc.edu.
 Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
 Series: IRAK Plate: 12 Row: n Column: 14
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.
 Location/Qualifiers
 1..816
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="Locustid:9616"
 /db_xref="taxon:9606"
 /clone="MGC:17274 IMAGE:4177613"
 /tissue type="Brain, anaplastic oligodendroglioma with 1p/19q loss"
 /clone_id="NCI CGAP_Brn67"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 37..378
 /codon_start=1

FEATURES
 SOURCE

CDS

/product="ring finger protein 7"
/protein_id="AA08627.1"
/db_xref="GI:14250389"
/translation="MADVDEGETCALASHSSGSGSGDKMFSLKKNNAVAMSWD
VECDTCAICRVQVMDACLRQAKENQEDCVVWGECHNSFHNCMSLWVKNNRCPLC
QODWVVRIGK"
BASE COUNT 226 a 178 c 213 g 199 t
ORIGIN

Query Match 25.2%; Score 32; DB 9; Length 816;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GCCATCTGCAGGTCAGGTGATGATGCCTG 45
Db 187 GCCATCTGCAGGTCAGGTGATGATGCCTG 218

RESULT 28
LOCUS BC005966 822 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, ring finger protein 7, clone MGC:14618 IMAGE:4069078,
mRNA, complete cds.
ACCESSION BC005966
VERSION BC005966.1 GI:13543635
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schultz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRVL Plate: 21 Row: 0 Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Genomescan gene
prediction. Similarity but not identity to protein.
FEATURES
source
1..822
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:9616"
/db_xref="taxon:9606"
/clone="MGC:14618 IMAGE:4069078"
/tissue_type="Kidney, hypernephroma"
/clone_lib="NIH_MGC_58"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
27..368
/codon_start=1
/product="ring finger protein 7"
/protein_id="AA08627.1"
/db_xref="GI:13543636"
/translation="MADVDEGETCALASHSSGSGSGDKMFSLKKNNAVAMSWD

VECDTCAICRVQVMDACLRQAKENQEDCVVWGECHNSFHNCMSLWVKNNRCPLC
QODWVVRIGK"
BASE COUNT 240 a 175 c 210 g 197 t
ORIGIN

Query Match 25.2%; Score 32; DB 9; Length 822;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GCCATCTGCAGGTCAGGTGATGATGCCTG 45
Db 177 GCCATCTGCAGGTCAGGTGATGATGCCTG 208

RESULT 29
LOCUS AF164679 836 bp mRNA linear PRI 14-OCT-1999
DEFINITION Homo sapiens ring finger protein CKBP1 mRNA, complete cds.
ACCESSION AF164679
VERSION AF164679.1 GI:5917673
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 836)
AUTHORS Son, M.-Y., Park, J.-W., Kim, Y.-S., Kang, S.-W., Marshak, D.R., Park, W. and
Bae, Y.-S.
TITLE Protein kinase CKII interacts with and phosphorylates the SAG
protein containing ring-H2 finger motif
JOURNAL Biochem. Biophys. Res. Commun. 263 (3), 743-748 (1999)
MEDLINE 99443734
PubMed 10512750
REFERENCE 2 (bases 1 to 836)
AUTHORS Son, M.-Y., Park, J.-W., Kim, Y.-S., Kang, S.-W., Marshak, D.R., Park, W.
and Bae, Y.-S.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1999) Department of Biochemistry, Kyungpook
National University, Taegu 702-701, Korea
FEATURES
source
1..836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
54..395
/note="SAG, contains an H2 ring finger motif"
/codon_start=1
/product="ring finger protein CKBP1"
/protein_id="AAD55984.1"
/db_xref="GI:5917674"
/translation="MADVDEGETCALASHSSGSGSGDKMFSLKKNNAVAMSWD
VECDTCAICRVQVMDACLRQAKENQEDCVVWGECHNSFHNCMSLWVKNNRCPLC
QODWVVRIGK"
BASE COUNT 228 a 185 c 223 g 200 t
ORIGIN

Query Match 25.2%; Score 32; DB 9; Length 836;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GCCATCTGCAGGTCAGGTGATGATGCCTG 45
Db 204 GCCATCTGCAGGTCAGGTGATGATGCCTG 235

RESULT 30
LOCUS AC118192 187353 bp DNA linear HTG 14-APR-2002
DEFINITION Mus musculus clone RP23-200C17, WORKING DRAFT SEQUENCE, 10 ordered
pieces.
ACCESSION AC118192
VERSION AC118192.1 GI:20147957
KEYWORDS HTG, HTGS_PHASE2, HTGS_DRAFT, HTGS_FULLTOP.

SOURCE ORGANISM	TITLE JOURNAL	REFERENCE AUTHORS
Mus musculus (house mouse)		
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
1 (bases 1 to 187353)		
Birtten,B., Linton,L., Nusbaum,C. and Lander,E.		
Mus musculus, clone RP23-200C17		
Unpublished		
2 (bases 1 to 187353)		
Birtten,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,L., Bloom,T., Boguslavsky,L., Booksgaller,B., Brown,A., Camarata,J., Campionio,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gairyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K., Lamazares,R., Landers,T., Lehoczeky,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Major,J., Margulis,N., Mathews,C., McCarthy,M., McEwan,P., McKernan,K., Meltrin,J., Meneses,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhthang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Riese,C., Rogov,P., Roman,J., Rossetti,M., Roy,A., Santos,R., Schauer,S., Schnepack,R., Semman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strassus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tophan,K., Travers,M., Travis,N., Trifillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wymann,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.		
Direct Submission		
Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
All repeats were identified using RepeatMasker:		
Smit, A.F.A. & Green, P. (1996-1997)		
http://ftp.genome.washington.edu/RM/RepeatMasker.html		
Genome Center		
Center: Whitehead Institute/ MIT Center for Genome Research		
Center code: WIBR		
Web site: http://www-seq.wi.mit.edu		
Contact: sequence_submissions@genome.wi.mit.edu		
Project Information		
Center project name: L19568		
Center clone name: 200 C_17		
Summary Statistics		
Sequencing vector: plasmid; n/a; 100% of reads		
Chemistry: Dye-terminator Big Dye; 100% of reads		
Assembly program: Phrap; version 0.960731		
Consensus quality: 184259 bases at least Q40		
Consensus quality: 185710 bases at least Q30		
Consensus quality: 186169 bases at least Q20		
Insert size: 194000; agarose-fp		
Insert size: 186453; sum-of-ctrls		
Quality coverage: 7.4 in Q20 bases; agarose-fp		
Quality coverage: 7.7 in Q20 bases; sum-of-ctrls		
NOTE: This is a 'working draft' sequence. It currently		
consists of 10 contigs. Gaps between the contigs		
are represented as runs of N. The order of the pieces		
is believed to be correct as given, however the sizes		
of the gaps between them are based on estimates that have		
provided by the submittor.		
This sequence will be replaced		
by the finished sequence as soon as it is available and		
the accession number will be preserved.		
1		
629: contig of 629 bp in length		
630		
729: gap of 100 bp		
730		
750: contig of 21 bp in length		
751		
850: gap of 100 bp		
851		
3212: contig of 2362 bp in length		
3213		
3312: gap of 100 bp		
3313		
6908: contig of 3596 bp in length		

FEATURES	SOURCE
* 6909	7008: gap of 100 bp
* 14434	config of 7426 bp in length
* 7009	14534: gap of 100 bp
* 14435	14534: gap of 100 bp
* 32181	config of 17647 bp in length
* 32182	32281: gap of 100 bp
* 32281	32281: gap of 100 bp
* 51167	config of 18886 bp in length
* 51168	51267: gap of 100 bp
* 51268	72870: config of 21603 bp in length
* 72871	72970: gap of 100 bp
* 149992	config of 77022 bp in length
* 149993	150092: gap of 100 bp
* 150093	187353: config of 37261 bp in length.
Location/Qualifiers	
1. 187353	
/organism="Mus musculus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10090"	
/clone="RP23-200C17"	
/clone_1fb="RP23-200C17"	
1. 629	
/note="assembly_fragment"	
730. 750	
/note="assembly_fragment"	
851. 3212	
/note="assembly_fragment"	
3313. 6908	
/note="assembly_fragment"	
7009. 14434	
/note="assembly_fragment"	
14535. 32181	
/note="assembly_fragment"	
32282. 51167	
/note="assembly_fragment"	
51268. 72870	
/note="assembly_fragment"	
72971. 149992	
/note="assembly_fragment"	
150093. 187353	
/note="assembly_fragment"	
clone_end:T7	
vector_side:right	
BASE COUNT	60628 a 34855 c 36129 g 54836 t 905 others
ORIGIN	
Query Match	25.2%; Score 32; DB 2; Length 187353;
Best Local Similarity	100.0%; Pred. No. 3,6e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Py	13 TGGCATCTGCAGGAGTCCAGGTGATGATGCTT 44
Db	46892 TGGCATCTGCAGGAGTCCAGGTGATGATGCTT 46923
RESULT 31	
AC095698/c	
LOCUS	AC095698 254832 bp DNA linear HTG 09-NOV-2002
DEFINITION	Rattus norvegicus clone CH230-9D13, WORKING DRAFT SEQUENCE.
ACCESSION	AC095698
KEYWORDS	AC095698.6 GI:24817898
SOURCE	HTG: HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
ORGANISM	Rattus norvegicus (Norway rat)
REFERENCE	
AUTHORS	1 (bases 1 to 254832)
	Muzny,D.,Marle, Metzker,M.,Lee, Abramzon,S., Adams,C., Alder,J.,
	Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angiano,D.,
	Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
	Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
	Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
	Bryant,N., Buhey,C., Burch,P., Burrill,K., Calderon,E.,
	Cadenas,V., Carter,K., Cavazos,I., Cessari,H., Cente,A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunatirne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Huily, S., Hume, J., Idelbird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpeth, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyar, C., Kowls, C., Kraft, C. L., Lebow, H., Lavan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensueta, L., Louieged, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokedumeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, U., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shervy, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Snead, A., Sodergren, E., Song, X. Z., Sorelle, R., Soes, J., Steidle, M., Strong, R., Sutton, A., Svacek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 254832)
Worley, K. C.

Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 254832)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23267899.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDBC
Center clone name: CH230-9D13
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 238918 bases at least Q40
Consensus quality: 241069 bases at least Q30
Consensus quality: 242304 bases at least Q20
Estimated insert size: 243228; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 254832: contig of 254832 bp in length.
Location/Qualifiers
1. 254832
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9D13"
1. 1482
/note="wgs end extension
clone_end:5p6"
2092. 2921
/note="clone_boundary
clone_end:5p6
site:ECORI
end sequence: BH339916"
23357. 235284
/note="wgs contig"
253718. 254832
/note="wgs contig"
complement(253801..254228)
/note="clone_boundary
clone_end:T7
site:ECORI
end sequence: BH339912"
BASE COUNT 63080 a 58350 c 57410 g 64047 t 11945 others
ORIGIN
Query Match 24.4%; Score 31; DB 2; Length 254832;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 CCATCGCAGGGTCCAGGATGATGATGCTG 45
Db 146354 CCATCGCAGGGTCCAGGATGATGATGCTG 146324
RESULT 32
BD096984
LOCUS BD096984 754 bp DNA linear PAT 27-AUG-2002
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096984.1 GI:22642572
VERSION BD096984.1 GI:22642572
KEYWORDS JP 2001526063-A/19.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 19 18-DEC-2001;

```

COMMENT      WARNER LAMBERT CO
              OS      Unidentified
              PN      JP 2001526063-A/19
              PD      18-DEC-2001
              PE      15-DEC-1998 JP 2000525451
              PR      19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
              YI SUN
PC      C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC      A61P39/06,
PC      A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
PC      C12N5/10,C12Q1/68,
PC      G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC      Topology: Linear;
CC      SAG:apoptosis sensitivity gene
FH      Key      Location/Qualifiers
FT      CDS      1..339
FT      mat_peptide 1..339.
FT      Location/Qualifiers
              1..754
              /organism="unidentified"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"
BASE COUNT   206 a 155 c 201 g 192 t
ORIGIN
Query Match      23.6%; Score 30; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 6.9e-07;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      14 GCCATCTGCAGGTCAGGTGATGATGCC 43
Db      151 GCCATCTGCAGGTCAGGTGATGATGCC 180
RESULT 33
LOCUS      BD096993      754 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION  BD096993.1 GI:22642581
VERSION     BD096993.1 GI:22642581
KEYWORDS    JP 2001526063-A/28.
SOURCE      unidentified
ORGANISM    unidentified
REFERENCE   1 (bases 1 to 754)
AUTHORS     Sun,Y.
TITLE       SAG:apoptosis sensitivity gene
PATENT      JP 2001526063-A 28 18-DEC-2001;
JOURNAL     WARNER LAMBERT CO
COMMENT     OS      Unidentified
              PN      JP 2001526063-A/28
              PD      18-DEC-2001
              PE      15-DEC-1998 JP 2000525451
              PR      19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
              YI SUN
PC      C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC      A61P39/06,
PC      A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
PC      C12N5/10,C12Q1/68,
PC      G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC      Topology: Linear;
CC      SAG:apoptosis sensitivity gene
FH      Key      Location/Qualifiers
FT      CDS      1..339
FT      mat_peptide 1..339.
FT      Location/Qualifiers
              1..754
              /organism="unidentified"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"
BASE COUNT   207 a 155 c 201 g 191 t

```

```

ORIGIN
Query Match      23.6%; Score 30; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 6.9e-07;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      14 GCCATCTGCAGGTCAGGTGATGATGCC 43
Db      151 GCCATCTGCAGGTCAGGTGATGATGCC 180
RESULT 34
LOCUS      AC119145      148290 bp      DNA      linear      HTG 11-JUN-2002
DEFINITION Rattus norvegicus chromosome 1 clone RP32-4G14, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
ACCESSION  AC119145
VERSION     AC119145.1 GI:20303534
KEYWORDS    HTG, HTGS_PHASE1, HTGS_FUL10P, HTGS_DRAFT.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
REFERENCE   1 (bases 1 to 148290)
AUTHORS     Taenzer,S., Monti,U., Gloeckner,G., Goesele,C., Baumgart,C.,
              Hubner,N. and Platzer,M.
TITLE       Rat chromosome 1 genomic sequence
JOURNAL     Unpublished
COMMENT     2 (bases 1 to 148290)
              Taenzer,S. and Platzer,M.
              Direct Submission
              Submitted (25-APR-2002) Genome Analysis, Institute of Molecular
              Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
              3 (bases 1 to 148290)
              Taenzer,S. and Platzer,M.
              Direct Submission
              Submitted (11-JUN-2002) Genome Analysis, Institute of Molecular
              Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
              ----- Genome Center
              Center: Institute of Molecular Biotechnology
              Center code: IMB
              Web site: http://genome.imb-jena.de/
              Contact: gscj-submit@genome.imb-jena.de
              ----- Project Information
              Center project name: RP32-4G14
              Center clone name: RP32-4G14
              ----- Summary Statistics
              Sequencing vector: pUC18; 100% of reads
              Chemistry: Dye-terminator Big Dye, 100% of reads
              Assembly program: Phrap; version 0.990329
              Consensus quality: 140730 bases at least Q40
              Consensus quality: 143676 bases at least Q30
              Consensus quality: 145595 bases at least Q20
              Quality coverage: 8.49
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

```

* be preserved.
* 1 25364: contig of 25364 bp in length
* 25365 25464: gap of unknown length
* 25466 38093: contig of 12629 bp in length
* 38094 38193: gap of unknown length
* 38194 83797: contig of 45604 bp in length
* 83798 83897: gap of unknown length
* 83898 90673: contig of 6776 bp in length
* 90674 90773: gap of unknown length
* 90774 107588: contig of 16815 bp in length
* 107589 107688: gap of unknown length
* 107689 112170: contig of 4482 bp in length
* 112171 112700: gap of unknown length
* 112701 119805: contig of 7535 bp in length
* 119806 119905: gap of unknown length
* 119906 137949: contig of 18044 bp in length
* 137950 138050: gap of unknown length
* 138051 146503: contig of 8454 bp in length
* 146504 146603: gap of unknown length
* 146604 148290: contig of 1687 bp in length.

FEATURES
    source
        1..148290
            location/Qualifiers
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /chromosome="1"
            /clone_lib="RPCT-33 BAC Library"
            /clone_id="31830 c 31870 g 41911 t 900 others"

BASE COUNT
    41779 a 31830 c 31870 g 41911 t 900 others

ORIGIN
    Query Match      23.6% Score 30; DB 2; Length 148290;
    Best Local Similarity 100.0%; Pred. No. 6.2e-07;
    Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 GATGTCAGCTGAAAAACAGCAGAGGACT 80
Db 17445 GATGTCAGCTGAAAAACAGCAGAGGACT 17474

RESULT 35
AC125667
LOCUS
DEFINITION
    Rattus norvegicus clone CH230-21C17, WORKING DRAFT SEQUENCE, 2
    unordered pieces.
ACCESSION
    AC125667
VERSION
    GI:30522767
KEYWORDS
    HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
SOURCE
    Rattus norvegicus (Norway rat)
ORGANISM
    Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
    Rattus.
    1 (bases 1 to 250425)
    Muzny,D.,Marle,, Metzker,M.,Lee,, Abramson,S., Adams,C., Alder,J.,
    Allen,C., Allen,H., Alshrocks,S., Amin,A., Anguiano,D.,
    Anyalebechi,V., Ayogaji,A., Ayodeji,M., Baca,E., Baden,H.,
    Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Bernhamed,F.,
    Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
    Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
    Cardenas,V., Carter,K., Cavazos,I., Ceesar,H., Center,A.,
    Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
    Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,K., D' Souza,L.,
    Davila,M.L., Davis,C., Davy-Catroll,L., De Anda,C., Dederich,D.,
    Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
    Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
    Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
    Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
    Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
    Gebregergis,E., Geer,K., Gill,R., Grady,A., Guerra,W., Guevara,W.,
    Guneratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
    Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
    Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hughes,M.,

```

TITLE

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Joliver,A., Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Komis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Louised,H., Lozada,R.J., Lu,X., Ma,J., Maheshwari,M., Mahidartine,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,G., Olampunsgoon,A., Pal,S., Parks,K., Nwakoelamoh,O., Okwou,G., Olampunsgoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plummer,F., Poidexter,A., Popovic,D., Primus,E., Pu,L., Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajd,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Swatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseca,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wiczysk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G., and Gibbs,R.A.

Direct Submission

2 (bases 1 to 250425).

Worley,K.C.

Direct Submission

Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 250425)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24819649. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GSKL

Center clone name: CH230-21C17

Summary Statistics

Assembly program: Atlas 3.0/

Consensus quality: 241733 bases at least Q40

Consensus quality: 244050 bases at least Q30

Estimated insert size: 249556; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

		NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)	
		* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.	
		* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	
		* 1 249202: contig of 249202 bp in length	
		* 249203 249302: gap of unknown length	
		* 249303 250425: contig of 1123 bp in length.	
FEATURES	source	location/Qualifiers	
		1..250425	
		/organism="Rattus norvegicus"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:10116"	
		/clone="CH230-21C17"	
		1..1392	
	misc_feature	/note="wgs_end extension clone_end:Sp6"	
		2873..3401	
	misc_feature	/note="clone boundary clone_end:Sp6 site:EcoRI"	
		end_sequence:BH268983"	
	misc_feature	245202..245446	
		/note="clone boundary clone_end:T7 site:EcoRI end_sequence:BH268982"	
	misc_feature	246080..247876	
		/note="wgs_end extension clone_end:T7"	
	misc_feature	247927..249202	
		/note="wgs_end extension clone_end:T7"	
BASE COUNT	70050 a	53545 c	53576 g 69593 t 3661 others
ORIGIN			
Query Match	23.6%; Score 30; DB 2; Length 250425;		
Best Local Similarity	100.0%; Pred. No. 6.2e-07;		
Matches	30; Conservative	0; Mismatches	0; Indels 0; Gaps 0
Oy	51 GATGTCAAGCTGAAAACAAGCAAGAGACT 80		
Db	106620 GATGTCAAGCTGAAAACAAGCAAGAGACT 106649		
RESULT 36			
LOCUS	AF312226	674 bp	mRNA linear PRI 06-SEP-2001
DEFINITION	Homo sapiens SAG splice variant mRNA, complete cds.		
ACCESSION	AF312226		
VERSION	AF312226.1	GI:13649605	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 674)		
AUTHORS	Duan,H., Wang,Y., Aviram,M., Swaroop,M., Luo,J.A., Bian,J., Tian,Y., Mueller,T., Bisgaler,C.L. and Sun,Y.		
TITLE	SAG, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents		
JOURNAL	Mol. Cell. Biol. 19 (4), 3145-3155 (1999)		
MEDLINE	99182502		
PubMed	10082581		
REFERENCE	2 (bases 1 to 674)		
AUTHORS	Swaroop,M., Gosink,M. and Sun,Y.		
TITLE	SAG/ROC2/Rbx2/Hrt2, a component of SCF E3 ubiquitin ligase: genomic structure, a splicing variant, and two family pseudogenes		

JOURNAL	DNA Cell Biol.	20 (7), 425-434	(2001)
MEDLINE	21398045		
PUBMED	11506706		
REFERENCE	3	(bases 1 to 674)	
AUTHORS	Sun, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-OCT-2000) Molecular Biology, Pfizer Global Research and Development, 2800 Plymouth Road, Ann Arbor, MI 48105, USA		
FEATURES	location/Qualifiers		
source	1..674		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="3"		
	/map="3q22-q24"		
	1..327		
CDS	/note="SAG-V; lacks zinc RING finger domain, has no activity as a component of SCF E3 ubiquitin ligase"		
	/codon_start=1		
	/product="SAG splice variant"		
	/protein_id="AAK37450.1"		
	/db_xref="GI:13649606"		
	/translation="MADVEDGEERFCALASHSGSSGSKGDKMPSLKKYNAVMMSPDVECTCAICRYQVWDEBGTGYRNMSEALNLINASBMGPDCRSGLTALVPVSVLASHOPCLLDHR"		
BASE COUNT	163 a	154 c	194 g 163 t
ORIGIN			
Query Match	22.8%; Score 29;	DB 9;	Length 674;
Best local similarity	100.0%;	Pred. NO. 2.9e-06;	
Matches	29;	Conservative 0;	Mismatches 0;
	Indels 0;	Gaps 0;	
OY	71 CAAGAGACTGTGTTGGTGTCGGAGAGA 99 		
Db	472 CAAAGAGACTGTGTTGGTGTCGGAGAGA 500		
RESULT 37			
BD096973		747 bp	DNA linear PAT 27-AUG-2002
LOCUS	BD096973		
DEFINITION	SAG:apoptosis sensitivity gene.		
ACCESSION	BD096973.1	GI:22642561	
VERSION	JP 2001526063-A/8.		
KEYWORDS	unidentified		
SOURCE	unclassified.		
ORGANISM	1 (bases 1 to 747)		
REFERENCE	Sun, Y.		
AUTHORS	SAG:apoptosis sensitivity gene		
TITLE	Patent: JP 2001526063-A 8 18-DEC-2001;		
JOURNAL	WARNER LAMBERT CO		
COMMENT	OS Unidentified		
	PN JP 2001526063-A/8		
	PD 18-DEC-2001		
	PF 15-DEC-1998 JP 2000525451		
	PR 19-DEC-1997 US 60/0661179, 11-SEP-1998 US 60/099840 PI		
	YI SUN		
	PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,		
	PC A61J39/06,		
	PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC		
	,C12N5/10,C12O1/68,		
	PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC		
	Strandedness: Double;		
	CC Topology: Linear;		
	CC SAG:apoptosis sensitivity gene		
	FH Key location/Qualifiers		
	FT CDS 1..270		
	FT mat_peptide 1..270.		
	FT location/Qualifiers		
	1..747		
FEATURES			
source	/organism="unidentified"		
	/mol_type="genomic DNA"		

BASE COUNT 204 a 155 c 197 g 191 t
ORIGIN

Query Match 22.8%; Score 29; DB 6; Length 747;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 CAAGAGACTGTGTGCTGGCGGAGA 99
Db 201 CAAGAGACTGTGTGCTGGCGGAGA 229

RESULT 38
BD096982 754 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096982.1 GI:22642570
VERSION
KEYWORDS JP 2001526063-A/17.
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 17 18-DEC-2001;
WARNER LAMBERT CO
COMMENT
OS Unidentified
PN JP 2001526063-A/17
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179, 11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P17/02, A61P35/00,
PC A61P39/06,
PC A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC
C12N5/10, C1201/68,
PC G01N33/50, G01N33/68, C12N15/00, A61K37/02, C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FEATURES
source 1..754
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t
ORIGIN

Query Match 22.8%; Score 29; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 CAAGAGACTGTGTGCTGGCGGAGA 99
Db 208 CAAGAGACTGTGTGCTGGCGGAGA 236

RESULT 39
BD096983 754 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096983.1 GI:22642571
VERSION
KEYWORDS JP 2001526063-A/18.
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 754)

AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 18 18-DEC-2001;
WARNER LAMBERT CO
COMMENT
OS Unidentified
PN JP 2001526063-A/18
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179, 11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P17/02, A61P35/00,
PC A61P39/06,
PC A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC
C12N5/10, C1201/68,
PC G01N33/50, G01N33/68, C12N15/00, A61K37/02, C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FEATURES
source 1..754
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 207 a 155 c 201 g 191 t
ORIGIN

Query Match 22.8%; Score 29; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 CAAGAGACTGTGTGCTGGCGGAGA 99
Db 208 CAAGAGACTGTGTGCTGGCGGAGA 236

RESULT 40
AC103666 112321 bp DNA linear HTG 20-DEC-2002
LOCUS
DEFINITION Mus musculus clone RP23-397H13, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC103666
VERSION
KEYWORDS AC103666.2 GI:27311471
HTG; HTGS_PHASE0.
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 112321)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-397H13
Unpublished
2 (bases 1 to 112321)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
Menus, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C.,
North, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (29-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 112321)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karakas, A., Kellis, C., Landers, T., Levine, R.,
Lindblad-Ton, K., Liu, G., MacLean, C., MacDonald, P., Major, J.,
Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupp, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (20-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 20, 2002 this sequence version replaced gi:17149550.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L19615
Center clone name: 397_H_13

* NOTE: This record contains 89 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 1147: contig of 1147 bp in length
* 1148 1247: gap of 100 bp
* 1248 2348: contig of 1101 bp in length
* 2349 2448: gap of 100 bp
* 2449 3511: contig of 1063 bp in length
* 3512 3611: gap of 100 bp
* 3612 4746: contig of 1135 bp in length
* 4747 4846: gap of 100 bp
* 4847 6030: contig of 1184 bp in length
* 6031 6130: gap of 100 bp
* 6131 7290: contig of 1160 bp in length
* 7291 7390: gap of 100 bp
* 7391 8577: contig of 1187 bp in length
* 8578 8678: gap of 100 bp
* 8679 9937: contig of 1260 bp in length
* 9938 10037: gap of 100 bp
* 10038 11192: contig of 1155 bp in length
* 11193 11292: gap of 100 bp
* 11293 12454: contig of 1162 bp in length
* 12455 12554: gap of 100 bp
* 12555 13718: contig of 1164 bp in length

13719 13818: gap of 100 bp
* 13819 15002: contig of 1184 bp in length
* 15003 15102: gap of 100 bp
* 15103 16234: contig of 1132 bp in length
* 16235 16334: gap of 100 bp
* 16335 17461: contig of 1127 bp in length
* 17462 17561: gap of 100 bp
* 17562 18650: contig of 1089 bp in length
* 18651 18750: gap of 100 bp
* 18751 20012: contig of 1162 bp in length
* 20013 21199: gap of 100 bp
* 21199 21992: contig of 1187 bp in length
* 21993 22100: gap of 100 bp
* 22100 22462: contig of 1163 bp in length
* 22463 22562: gap of 100 bp
* 22563 23720: contig of 1158 bp in length
* 23721 23820: gap of 100 bp
* 23821 25005: contig of 1185 bp in length
* 25006 25105: gap of 100 bp
* 25106 26265: contig of 1160 bp in length
* 26266 26365: gap of 100 bp
* 26366 27528: contig of 1163 bp in length
* 27529 27628: gap of 100 bp
* 27629 28788: contig of 1160 bp in length
* 28789 28888: gap of 100 bp
* 28889 30034: contig of 1146 bp in length
* 30035 30134: gap of 100 bp
* 30135 31236: contig of 1102 bp in length
* 31237 31336: gap of 100 bp
* 31337 32513: contig of 1176 bp in length
* 32514 32612: gap of 100 bp
* 32613 33813: contig of 1201 bp in length
* 33814 33914: gap of 100 bp
* 33915 35107: contig of 1194 bp in length
* 35108 35207: gap of 100 bp
* 35209 36431: contig of 1224 bp in length
* 36432 36531: gap of 100 bp
* 36532 37701: contig of 1170 bp in length
* 37702 37801: gap of 100 bp
* 37802 38967: contig of 1166 bp in length
* 38968 39067: gap of 100 bp
* 39068 40235: contig of 1168 bp in length
* 40236 40335: gap of 100 bp
* 40336 41492: contig of 1157 bp in length
* 41493 41592: gap of 100 bp
* 41593 42691: contig of 1099 bp in length
* 42692 42791: gap of 100 bp
* 42792 43970: contig of 1179 bp in length
* 43971 44070: gap of 100 bp
* 44071 45150: contig of 1080 bp in length
* 45151 45250: gap of 100 bp
* 45251 46347: contig of 1097 bp in length
* 46348 46447: gap of 100 bp
* 46449 47608: contig of 1161 bp in length
* 47609 47708: gap of 100 bp
* 47709 48881: contig of 1173 bp in length
* 48882 48981: gap of 100 bp
* 48982 50240: contig of 1259 bp in length
* 50241 50340: gap of 100 bp
* 50341 51544: contig of 1204 bp in length
* 51545 51644: gap of 100 bp
* 51645 52828: contig of 1184 bp in length
* 52829 52928: gap of 100 bp
* 52929 54157: contig of 1229 bp in length
* 54158 54257: gap of 100 bp
* 54258 55450: contig of 1199 bp in length
* 55451 55550: gap of 100 bp
* 55551 56731: contig of 1181 bp in length
* 56732 56831: gap of 100 bp
* 56832 58019: contig of 1188 bp in length
* 58020 58119: gap of 100 bp
* 58120 59285: contig of 1166 bp in length
* 59286 59385: gap of 100 bp

Cleveland, C.D., Cox, C., Coyle, M.D., Bathone, S.R., David, R.,
David, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Deen, A.L., Ding, Y., Dinik, H.H.,
Douthwaite, K.U., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Einhart, C., Edgar, D., Edwards, C.C., Elhat, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabsi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,
Hollins, B., Homsl, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Kovach, J.,
Kovar, C., Kratovic, J., Kurehi, A., Landry, N., Leal, B., Lewis, I.C.,
Lewis, L., Li, U., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Lousaged, H., Lozado, R.U., Lu, X., Luckier, A., Luckier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,
Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, W., Neal, D.,
Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N.,
Nickerson, E., Nwokkenwo, S., Ogulu, M., Okunolu, G., Oragune, N.,
Ovidio, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L.,
Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M.,
Rojs, A., Rojibokan, I., Rolle, M., Ruiz, S., Savey, G., Scherer, S.,
Scott, G., Shen, H., Shoehrati, N., Sisson, I., Sodergren, E.,
Soniak, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,
Tabori, P., Tamertisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C.,
Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L.,
Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,
Warren, R., Washington, C., Watlington, S., Williams, G.,
Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,
Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and
Gibbs, R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 35638)

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (24-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 35638)

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (25-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 35638)

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 5 (bases 1 to 35638)

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (28-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Jul 30, 2002 this sequence version replaced gi|20976460.
INFORMATION: <http://www.hgscc.dcm.tmc.edu> or email gc-help@bcm.tmc.edu

ANNOTATION OF FEATURES:
STSS are identified using ePCR (Genome Res. 7:541-550) searches

```

a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeat are identified using RepeatMasker (A. Smit and P. Green
unpublished) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.
SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.
QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.
FEATURES
source
location/Qualifiers
1..35638
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-144C9"
1..1895
/note="overlaps bases 56107..58001 of clone ACO92978"
/function="clone overlap"
14..341
/standard_name="SHGC-105132"
complement(1283..1580)
/rpt_family="AluSx"
1618..1906
/rpt_family="AluSx"
1907..1986
/rpt_family="AluSx"
complement(2109..2203)
/rpt_family="L1M4"
complement(2204..2463)
/rpt_family="AluSx"
complement(2464..2700)
/rpt_family="L1M4"
complement(2701..2819)
/rpt_family="AluSc"
complement(2843..3539)
/rpt_family="L1M4"
complement(3541..3642)
/rpt_family="AluSg/x"
3643..3758
/rpt_family="FLM_A"
complement(3777..4993)
/rpt_family="L1MEC"
complement(4994..5287)
/rpt_family="AluSg"
complement(5288..5404)
/rpt_family="L1MEC"
complement(5405..5716)
/rpt_family="AluSx"
complement(5717..6153)
/rpt_family="L1MEC"
complement(6154..6379)
/rpt_family="AluSg/x"
complement(6380..6394)
/rpt_family="L1MEC"
complement(6427..6570)
/rpt_family="MER91A"

```

```

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
  http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

URES
source
  location/Qualifiers
  1..35638
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="3"
  /clone="RP11-144C9"
  1..1895
  /note="overlaps bases 56107..58001 of clone AC092978"
  /function="clone overlap"
  14..341
  /standard_name="SHGC-105132"
  complement(1283..1580)
  /rpt_family="AluX*"
  1618..1906
  /rpt_family="AluX*"
  1907..1986
  /rpt_family="(GGGGA)n"
  complement(2109..2203)
  /rpt_family="L1M4"
  complement(2204..2463)
  /rpt_family="AluX*"
  complement(2464..2700)
  /rpt_family="L1M4"
  complement(2701..2819)
  /rpt_family="AluSc"
  complement(2843..3359)
  /rpt_family="L1M4"
  complement(3541..3642)
  /rpt_family="AluSg/x"
  3643..3758
  /rpt_family="FLAM A"
  complement(3777..4493)
  /rpt_family="L1MBC"
  complement(4994..5287)
  /rpt_family="AluSg"
  complement(5288..5404)
  /rpt_family="L1MBC"
  complement(5405..5716)
  /rpt_family="AluX*"
  complement(5717..6153)
  /rpt_family="L1MBC"
  complement(6154..6379)
  /rpt_family="AluSg/x"
  complement(6380..6394)
  /rpt_family="L1MBC"
  complement(6427..6570)
  /rpt_family="MER91A"

```

```

repeat_region      7418..7561
                    /rpt_family="AluSg/x"
repeat_region      8094..8139
                    /rpt_family="(TG)n"
repeat_region      complement(10186..10802)
                    /rpt_family="L2"
repeat_region      11949..12489
                    /rpt_family="HML1"
repeat_region      12490..12772
                    /rpt_family="AluSx"
repeat_region      12773..12804
                    /rpt_family="(TAAAA)n"
repeat_region      12805..12820
                    /rpt_family="AluSx"
repeat_region      12821..12896
                    /rpt_family="HML1"
repeat_region      12989..13063
                    /rpt_family="AluSg/x"
repeat_region      13069..13452
                    /rpt_family="HML1"
repeat_region      13453..13585
                    /rpt_family="FLAM_C"
repeat_region      13586..13609
                    /rpt_family="HML1"
repeat_region      13619..13654
                    /rpt_family="G-rich"
repeat_region      13653..13809
                    /rpt_family="G-rich"
repeat_region      complement(13819..14048)
                    /rpt_family="AluSg"
repeat_region      14077..14170
                    /rpt_family="MER113"
repeat_region      complement(14655..14830)
                    /rpt_family="MUT1U"
repeat_region      complement(14852..15139)
                    /rpt_family="AluSg"

Query Match      19.7% Score 25; DB 9; Length 35638;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      14 GCCATCTGCAGGCTCCAGGTGATGG 38
Db      25655 GCCATCTGCAGGCTCCAGGTGATGG 25679

RESULT 44
AC106176/c      238330 bp      DNA      linear      HTG 13-MAY-2003
LOCUS      Rattus norvegicus clone CH230-20P16, WORKING DRAFT SEQUENCE, 3
DEFINITION      unsorted pieces.
ACCESSION      AC106176
VERSION      AC106176.5 GI:30578781
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 238330)
Muzny,D,Marie, Metzker,M, Lee, Adamson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alspbrooks,S, Amin,A, Anguiano,D,
Anyalelechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Biswal,N, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Cesar,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, P'Souza,L,
Davila,M,L, Davis,C, Davy-Carrillo,L, De Anda,C, Dedertich,D,
Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,C, Divya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,

```

```

TITLE      JOURNAL
REFERENCE      Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
AUTHORS      Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
TITLE      Gebregeorgis,B, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,M,
JOURNAL      Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K,
TITLE      Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
AUTHORS      Hernandez,R, Hines,S, Hladun,S,L, Hodson,A, Hogues,M,
TITLE      Hollins,B, Howells,S, Huylk,S, Hune,J, Idlibird,D, Jackson,A,
JOURNAL      Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
TITLE      Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,
AUTHORS      Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
TITLE      Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
JOURNAL      Lorensushewa,L, Loulseged,H, Lozado,R,J, Lu,X, Ma,J,
TITLE      Maheshwari,M, Mahindarane,M, Mahmoud,M, Malloy,K, Mangum,A,
AUTHORS      Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
TITLE      Manning,S, McLeod,M,P, McNeill,T,Z, Meenen,B,
JOURNAL      Mliostavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
TITLE      Morgan,M, Morris,K, Morris,S, Mundaas,M, Murphy,M, Natr,L,
AUTHORS      Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,
TITLE      Nockelemech,O, Okwomu,G, Olarunsaagoo,A, Pal,S, Parks,K,
JOURNAL      Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C,
TITLE      Plopper,F, Poindester,A, Popovic,D, Primus,E, Fu,J,
AUTHORS      Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
TITLE      Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
JOURNAL      Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,
TITLE      Sanders,M, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H,
AUTHORS      Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajic,D,
TITLE      Sneed,A, Sodergren,B, Song,X,Z, Sorelle,R, Sosa,J,
JOURNAL      Steimle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C,
TITLE      Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K,
AUTHORS      Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J,
TITLE      Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F,
JOURNAL      Williams,G, Willson,R, Wiczcyk,R, Wooden,H, Worley,K,
TITLE      Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V,
AUTHORS      Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von
TITLE      Niederhaubeern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O,
JOURNAL      Weinstein,G, and Gibbs,R.A.
REFERENCE      Direct Submission
AUTHORS      Unpublished
TITLE      2 (bases 1 to 238330)
JOURNAL      Worley,K.C.
REFERENCE      Direct Submission
AUTHORS      Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
TITLE      of Molecular and Human Genetics, Baylor College of Medicine, One
JOURNAL      Baylor Plaza, Houston, TX 77030, USA
REFERENCE      3 (bases 1 to 238330)
AUTHORS      Rat Genome Sequencing Consortium.
TITLE      Direct Submission
JOURNAL      Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
REFERENCE      of Molecular and Human Genetics, Baylor College of Medicine, One
AUTHORS      Baylor Plaza, Houston, TX 77030, USA
TITLE      COMMENT
JOURNAL

```

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both ends sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GUOH
 Center clone name: CH230-20P16
 ----- Summary Statistics
 Assembly program: Atlas 3.0;

	* NOTE: Estimated insert size may differ from sequence length
	(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
	* NOTE: This is a 'working draft' sequence. It currently
	* consists of 3 contigs. The true order of the pieces
	* is not known and their order in this sequence record is
	* arbitrary. Gaps between the contigs are represented as
	* runs of N, but the exact sizes of the gaps are unknown.
	* This record will be updated with the finished sequence
	* as soon as it is available and the accession number will
	* be preserved.
	*
	* 1 236056: contig of 236056 bp in length
	* 236057 236156: gap of unknown length
	* 236157 237213: contig of 1057 bp in length
	* 237214 237313: gap of unknown length
	* 237314 238330: contig of 1017 bp in length.
	*
FEATURES	Location/Qualifiers
source	1..238330
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10116"
	/clone="CH230-20P16"
	1..1582
misc_feature	/note="wgs contig"
BASE COUNT	64585 a 53668 c 53820 g 62326 t 3931 others
ORIGIN	
Query Match	19.7%; Score 25; DB 2; Length 238330;
Best Local Similarity	100.0%; Pred. No. 0.00078;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	14 GCCATCTGCAGGTCAGGTGATGG 38
Dd	164432 GCCATCTCAGGTCAGGTGATGG 164408
RESULT 45	
AC0098496/c	
LOCUS	Rattus norvegicus clone CH230-23C1, WORKING DRAFT SEQUENCE, 4
DEFINITION	AC0098496 245476 bp DNA linear HTG 10-MAY-2003
ACCESSION	unordered pieces.
KEYWORDS	AC0098496 GI:30521845
SOURCE	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
ORGANISM	Rattus norvegicus (Norway rat)
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1 (bases 1 to 245476) Munzy,D.Marie., Metzger,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angilano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalto,K., Blair,J., Blankenhorn,K., Blyth,P., Brown,M., Bryant,N., Bhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Claveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., pSouza,L., Devila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Deramo,C., Ding,X., Dim,H., Divya,K., Diaper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Georgopoulos,E., Geor,K., Gill,R., Grady,M., Guerra,W., Guerrero,W., Gunaratne,P., Haaland,P., Hamli,C., Hamilton,C., Hamilton,K., Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J.,

Hemandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Hollins,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovac,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorenshaw,L., Louised,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartine,M., Mahmud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 McWhirley,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Natr,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwaokiemeh,O., Okwomu,G., Olarunsgoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plappier,F., Polndexter,A., Popovic,D., Primus,E., Pu,L.,
 Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Rellly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojars,A., Rose,M., Rose,R., Rul,S.,J.,
 Sanders,W., Savary,G., Scheier,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Snaigs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steinhle,M., Strong,R., Sutton,A., Swatck,A., Taber,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,P., Usmani,K.,
 Valas,R., Vera,V., Villaseña,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 245476)
 Worley,K.C.
 Direct Submission
 Submitted (24-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 245476)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 10, 2003 this sequence version replaced gi:24819573.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/Projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 ----- Project name: GHEY
 Center project name: CH230-23C21
 Center clone name: CH230-23C21
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 228381 bases at least Q40
 Consensus quality: 230024 bases at least Q30
 Consensus quality: 231602 bases at least Q20
 Estimated insert size: 236545; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation


```

*****
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 5516: contig of 5516 bp in length
* 5517 5616: gap of unknown length
* 5617 242823: contig of 237207 bp in length
* 242824 243958: gap of unknown length
* 243959 244058: contig of 1035 bp in length
* 244059 245476: gap of unknown length
* 245476: contig of 1418 bp in length.
* Location/Qualifiers
*     1. 245476
*        /organism="Rattus norvegicus"
*        /mol_type="genomic DNA"
*        /db_xref="taxon:10116"
*        /clone="CH230-23C21"
*        1. 1560
*        /note="wgs_end_extension"
*        clone_end:17
*        complement(3889..4558)
*        /note="clone_boundary"
*        clone_end:17
*        site:ECORI
*        end sequence: BH361779"
* 5617..6729
*        /note="wgs_contig"
*        69231..103385
*        /note="clone_boundary"
*        clone_end:Sp6
*        site:ECORI
*        end sequence: BH361780"
*
BASE COUNT  67659 a 50656 c 50781 g 63728 t 12652 others
ORIGIN
Query Match      19.7%; Score 25; DB 2; Length 245476;
Best Local Similarity 100.0%; Pred. No. 0.00077;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 GCCATCTGCAGGTCGAGTGATGG 38
Db 69519 GCCATCTGCAGGTCGAGTGATGG 69495
*****
RESULT 46
AC118520/c 189541 bp DNA linear HTG 20-NOV-2002
LOCUS      Rattus norvegicus clone CH230-304J21, *** SEQUENCING IN PROGRESS
DEFINITION *** 2 unordered pieces.
ACCESSION  AC118520
VERSION     AC118520.4 GI:25138114
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 189541)
REFERENCE   Muzny, D., Marlet, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
AUTHORS     Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
            Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
            Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
            Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
            Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
            Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
            Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

```

```

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgescu, E., Geer, K., Gill, R., Grady, M., Guert, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Huylk, S., Hume, J., Idler, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Liu, J.,
Kowis, C., Kratt, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Louised, H., Lozada, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
Nankervyls, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackemele, O., Okunolu, G., Olarnpunasoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puato, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.M., Reich, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, J., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Woley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yan, J., Yoon, L., Yoon, V.,
Yu, P., Zhang, Y., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 189541)
REFERENCE   Worley, K.C.
AUTHORS     Direct Submission
TITLE       Submitted (18-APR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
JOURNAL
3 (bases 1 to 189541)
REFERENCE   Rat Genome Sequencing Consortium.
AUTHORS     Direct Submission
TITLE       Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
JOURNAL
COMMENT     On Nov 20, 2002 this sequence version replaced gi:33914484.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
*****
Center: Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

```



```

----- Project Information
Center project name: GUNO
Center clone name: CH230-304J21
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 173827 bases at least Q40
Consensus quality: 175964 bases at least Q30
Consensus quality: 176952 bases at least Q20
Estimated insert size: 178045; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 188414: contig of 188414 bp in length
* 188415 188514: gap of unknown length
* 188515 189541: contig of 1027 bp in length.
Location/Qualifiers
1.189541
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-304J21"
121774..123263
misc_feature
/note="wgs_contig"
complement(184170..185020)
/misc_feature
/note="clone boundary"
clone_end:T7
site:
end sequence:B225638"
187321..188414
/misc_feature
/note="wgs_end_extension"
clone="wgs_end:T7"

BASE COUNT 49137 a 39834 c 38226 g 50535 t 11809 others
ORIGIN

Query Match 18.9%; Score 24; DB 2; Length 189541;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 GGCCTGCTTGTGCTCGGCGAGA 99
|||||
Db 93758 GGCCTGCTTGTGCTCGGCGAGA 93755

RESULT 47
AC139954/c 208141 bp DNA linear HTG 27-MAR-2003
LOCUS Rattus norvegicus clone CH230-90M19, WORKING DRAFT SEQUENCE, 52
DEFINITION
unordered pieces.
AC139954
AC139954
AC139954
AC139954.3 GI:28626559
VERSION HTG: HTGS PHASE1; HTGS DRAFT.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 208141)
REFERENCE
Wuzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Aisbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benhmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blych,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
JOURNAL
AUTHORS
REFERENCE
AUTHORS
COMMENT

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
David,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Georgievski,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
Gunnarsson,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howell,S., Hulky,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Kapathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Koyat,C.,
Kovis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,
Miner,G., Munja,E., Montemayor,D., Moore,S., Morgan,M., Morris,K.,
Morris,S., Mundasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
Newton,N., Nguyen,N., Norris,S., Mwoketeme,O., Okunolu,G.,
Olatunbosun,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,
Perez,A., Perez,L., Pfankuch,C., Plopper,F., Poindexter,A.,
Popovic,D., Primus,E., Pu,L.L., Puazo,M., Quiroz,J., Rachin,B.,
Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
Reuter,M., Richards,S., Riggs,F., Riggs,C., Rodkey,T., Rojas,A.,
Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S.,
Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
Sison,I., Sitter,C.D., Smajs,D., Speed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,
Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villanada,D.,
Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczek,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 208141)
Worley,K.C.
Submitted (17-FEB-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 208141)
Worley,K.C.
Direct Submission
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 1, 2003 this sequence version replaced gi:28603887.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KDOM
Center clone name: CH230-90M19
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 178143 bases at least Q40
Consensus quality: 184319 bases at least Q30
Consensus quality: 188864 bases at least Q20
Estimated insert size: 182136; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

```

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.bjsc.bcm.tmc.edu/docs/genbank/draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 52 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of 'N', but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 163: contig of 1163 bp in length
 * 1164 1263: gap of unknown length
 * 1264 2611: contig of 1348 bp in length
 * 2612 2711: gap of unknown length
 * 2712 3836: contig of 1125 bp in length
 * 3837 3936: gap of unknown length
 * 3937 5317: contig of 1381 bp in length
 * 5318 5417: gap of unknown length
 * 5418 6568: contig of 1151 bp in length
 * 6569 7872: gap of unknown length
 * 7873 7972: gap of unknown length
 * 7973 9049: contig of 1077 bp in length
 * 9050 9149: gap of unknown length
 * 9150 10503: contig of 1354 bp in length
 * 10504 10603: gap of unknown length
 * 10604 11740: contig of 1137 bp in length
 * 11741 11840: gap of unknown length
 * 11841 12996: contig of 1156 bp in length
 * 12997 13096: gap of unknown length
 * 13097 14621: contig of 1525 bp in length
 * 14622 14721: gap of unknown length
 * 14722 15860: contig of 1139 bp in length
 * 15861 15960: gap of unknown length
 * 15961 17457: contig of 1497 bp in length
 * 17458 17557: gap of unknown length
 * 17558 19168: contig of 1611 bp in length
 * 19169 19268: gap of unknown length
 * 19269 20749: contig of 1481 bp in length
 * 20750 20849: gap of unknown length
 * 20850 22819: contig of 1970 bp in length
 * 22820 22919: gap of unknown length
 * 22920 25192: contig of 2273 bp in length
 * 25193 25292: gap of unknown length
 * 25293 27234: contig of 1942 bp in length
 * 27235 27334: gap of unknown length
 * 27335 28926: contig of 1592 bp in length
 * 28927 29026: gap of unknown length
 * 29027 30173: contig of 1147 bp in length
 * 30174 30273: gap of unknown length
 * 30274 32443: contig of 2170 bp in length
 * 32444 32543: gap of unknown length
 * 32544 34092: contig of 1549 bp in length
 * 34093 34192: gap of unknown length
 * 34193 37288: contig of 3096 bp in length
 * 37289 37388: gap of unknown length
 * 37389 40619: contig of 3231 bp in length
 * 40620 40719: gap of unknown length
 * 40720 43147: contig of 2428 bp in length
 * 43148 43247: gap of unknown length
 * 43248 45341: contig of 2094 bp in length
 * 45342 45441: gap of unknown length
 * 45442 49153: contig of 3712 bp in length
 * 49154 49253: gap of unknown length
 * 49254 52591: contig of 3338 bp in length
 * 52592 52691: gap of unknown length
 * 52692 55092: contig of 2401 bp in length
 * 55093 55192: gap of unknown length
 * 55193 56774: contig of 4482 bp in length
 * 56775 59774: gap of unknown length
 * 59775 62517: contig of 2743 bp in length
 * 62518 62617: gap of unknown length

* 62618 65416: contig of 2799 bp in length
 * 65417 65516: gap of unknown length
 * 65517 65517: contig of 3848 bp in length
 * 65965 69464: gap of unknown length
 * 69465 73440: contig of 3976 bp in length
 * 73441 73540: gap of unknown length
 * 73541 77575: contig of 4035 bp in length
 * 77576 77675: gap of unknown length
 * 77676 82110: contig of 4435 bp in length
 * 82111 82210: gap of unknown length
 * 82211 87479: contig of 5269 bp in length
 * 87480 87579: gap of unknown length
 * 87580 93557: contig of 5978 bp in length
 * 93558 93658: gap of unknown length
 * 93659 98354: contig of 4597 bp in length
 * 98355 104083: gap of unknown length
 * 104084 104183: gap of unknown length
 * 111520 111520: contig of 7337 bp in length
 * 111521 111620: gap of unknown length
 * 111621 116695: contig of 5075 bp in length
 * 116696 116795: gap of unknown length
 * 116796 123174: contig of 6379 bp in length
 * 123175 123274: gap of unknown length
 * 123275 131678: contig of 8404 bp in length
 * 131679 131778: gap of unknown length
 * 131779 138679: contig of 6901 bp in length
 * 138680 138779: gap of unknown length
 * 138780 146691: contig of 7912 bp in length
 * 146692 146791: gap of unknown length
 * 146792 153400: contig of 6609 bp in length
 * 153401 153500: gap of unknown length
 * 153501 161119: contig of 7619 bp in length
 * 161120 161219: gap of unknown length
 * 161220 171756: contig of 10537 bp in length

Query Match

Best Local Similarity 100.0%; Pred. No. 0.0032;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

76 GGACTGTGTTGCTGCTGGGAGA 99

Db

12567 GGACTGTGTTGCTGCTGGGAGA 12544

RESULT 48

AC094582/c

LOCUS

Rattus norvegicus clone CH230-4M15, WORKING DRAFT SEQUENCE.

AC094582.7

DEFINITION

HTG_09-MAY-2003

AC094582.7

KEYWORDS

HTG_09-MAY-2003

AC094582.7

SOURCE

Rattus norvegicus

AC094582.7

ORGANISM

Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;

AC094582.7

REFERENCE

1 (bases 1 to 243487)

Muzny, D., Marz, M., Metker, M., Lee, A., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amn, A., Angiano, D., Anyalelechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandarraia, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blaudenburg, K., Blyth, P., Brown, M., Bryant, N., Burch, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, X., Dim, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Georgiadis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guetara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hu, Y. S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louised, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindarcne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mabin, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackemeah, O., Okmonu, G., Olarnpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L. L., Pua, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wiczek, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Unpublished
2 (bases 1 to 243487)
Worley, K. C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 243487)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:2481567.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAYK
Center clone name: CH230-4M15
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 216410 bases at least Q40
Consensus quality: 218103 bases at least Q30

Consensus quality: 219151 bases at least Q20
Estimated insert size: 224236; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 243487: contig of 243487 bp in length.
Location/Qualifiers
1. 243487
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-4M15"
1. 2254
/note="wgs end extension
complement(3914..4636)
/note="clone_boundary
clone_end:17
site:ECORI
end_sequence:BH309319"
235713..236568
/note="clone boundary
clone_end:Sp6
site:ECORI
end_sequence:BH309321"
242181..243487
/note="wgs end extension
clone_end:Sp6"
misc_feature
62126 a 49267 c 47601 g 61044 t 23449 others
BASE COUNT
ORIGIN
Query Match 18.9%; Score 24; DB 2; Length 243487;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 GCACTGCTGTGCTGCTGGGAGCA 99
DB 17733 GCACTGCTGTGCTGCTGGGAGCA 17710

RESULT 49
AC104983 59155 bp DNA linear HTG 22-DEC-2001
LOCUS
DEFINITION Homo sapiens chromosome 18 clone RP11-352C3 map 18, LOW-PASS
SEQUENCE SAMPLING.
AC104983
VERSION AC104983.1 GI:17977361
KEYWORDS HTG; HTGS; PHASEO.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 59155)
Birtten, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 18, clone RP11-352C3
Unpublished
2 (bases 1 to 59155)
Birtten, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choe, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde,S., Gord,S., Coyette,M., Graham,L., Grand-Pierre,N.,
Jones,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jagos,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazaras,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Menus,L., Milova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Romanov,I., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnpack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Struss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zaitoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: L22883
Center clone name: 352_C_3

* NOTE: This record contains 70 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1	724:	contig of 724 bp	in length
725	824:	gap of 100 bp	
825	1582:	contig of 758 bp	in length
1583	1682:	gap of 100 bp	
1683	2409:	contig of 727 bp	in length
2410	2509:	gap of 100 bp	
2510	3235:	contig of 726 bp	in length
3236	3335:	gap of 100 bp	
3336	4090:	contig of 755 bp	in length
4091	4190:	gap of 100 bp	
4191	4955:	contig of 765 bp	in length
4956	5055:	gap of 100 bp	
5056	5780:	contig of 725 bp	in length
5781	5880:	gap of 100 bp	
5881	6643:	contig of 763 bp	in length
6644	6743:	gap of 100 bp	
6744	7496:	contig of 753 bp	in length
7497	7596:	gap of 100 bp	
7597	8348:	contig of 752 bp	in length
8349	8448:	gap of 100 bp	
8449	9217:	contig of 769 bp	in length
9218	9317:	gap of 100 bp	
9318	10093:	contig of 776 bp	in length
10094	10193:	gap of 100 bp	
10194	10924:	contig of 731 bp	in length
10925	11024:	gap of 100 bp	
11025	11793:	contig of 769 bp	in length
11793	11893:	gap of 100 bp	
11794	12661:	contig of 768 bp	in length
11894	12761:	gap of 100 bp	
12662	13525:	contig of 764 bp	in length
12762			

13526	13625:	gap of 100 bp	
13626	14347:	contig of 722 bp	in length
13627	14447:	gap of 100 bp	
13628	14448:	contig of 725 bp	in length
13629	15172:	contig of 725 bp	in length
13630	15173:	gap of 100 bp	
13631	15272:	contig of 735 bp	in length
13632	16007:	contig of 732 bp	in length
13633	16107:	gap of 100 bp	
13634	16108:	contig of 748 bp	in length
13635	16855:	contig of 748 bp	in length
13636	16955:	gap of 100 bp	
13637	16956:	contig of 756 bp	in length
13638	17711:	gap of 100 bp	
13639	17712:	contig of 751 bp	in length
13640	18568:	contig of 751 bp	in length
13641	18569:	gap of 100 bp	
13642	18669:	contig of 753 bp	in length
13643	19421:	contig of 753 bp	in length
13644	19521:	gap of 100 bp	
13645	20229:	contig of 708 bp	in length
13646	20329:	gap of 100 bp	
13647	20330:	contig of 732 bp	in length
13648	21061:	contig of 732 bp	in length
13649	21062:	gap of 100 bp	
13650	21161:	contig of 748 bp	in length
13651	22009:	gap of 100 bp	
13652	22010:	contig of 764 bp	in length
13653	22773:	gap of 100 bp	
13654	22774:	contig of 750 bp	in length
13655	23623:	contig of 750 bp	in length
13656	23723:	gap of 100 bp	
13657	23724:	contig of 766 bp	in length
13658	24489:	contig of 766 bp	in length
13659	24590:	gap of 100 bp	
13660	25354:	contig of 765 bp	in length
13661	25454:	gap of 100 bp	
13662	26174:	contig of 720 bp	in length
13663	26274:	gap of 100 bp	
13664	26275:	contig of 769 bp	in length
13665	27044:	gap of 100 bp	
13666	27144:	contig of 757 bp	in length
13667	27901:	gap of 100 bp	
13668	28000:	contig of 753 bp	in length
13669	28753:	contig of 753 bp	in length
13670	28853:	gap of 100 bp	
13671	28854:	contig of 719 bp	in length
13672	29572:	contig of 719 bp	in length
13673	29672:	gap of 100 bp	
13674	30419:	contig of 747 bp	in length
13675	30519:	gap of 100 bp	
13676	31262:	contig of 743 bp	in length
13677	31362:	gap of 100 bp	
13678	32052:	contig of 690 bp	in length
13679	32152:	gap of 100 bp	
13680	32909:	contig of 757 bp	in length
13681	33009:	gap of 100 bp	
13682	33736:	contig of 727 bp	in length
13683	33836:	gap of 100 bp	
13684	34598:	contig of 762 bp	in length
13685	34698:	gap of 100 bp	
13686	35456:	contig of 758 bp	in length
13687	35556:	gap of 100 bp	
13688	36274:	contig of 718 bp	in length
13689	36374:	gap of 100 bp	
13690	37111:	contig of 737 bp	in length
13691	37211:	gap of 100 bp	
13692	37954:	contig of 743 bp	in length
13693	38054:	gap of 100 bp	
13694	38783:	contig of 729 bp	in length
13695	38883:	gap of 100 bp	
13696	39686:	contig of 803 bp	in length
13697	39786:	gap of 100 bp	
13698	40535:	contig of 749 bp	in length
13699	40635:	gap of 100 bp	
13700	41387:	contig of 752 bp	in length
13701	41487:	gap of 100 bp	
13702	42254:	contig of 767 bp	in length
13703	42354:	gap of 100 bp	
13704	43046:	contig of 632 bp	in length
13705	43146:	gap of 100 bp	
13706	43885:	contig of 739 bp	in length
13707	43985:	gap of 100 bp	

```

* 43986 44741: contig of 756 bp in length
* 44742 44841: gap of 100 bp
* 44842 44842: contig of 763 bp in length
* 44842 45604: contig of 763 bp in length
* 45605 45704: gap of 100 bp
* 45705 46429: contig of 725 bp in length
* 46430 46529: gap of 100 bp
* 46530 47244: contig of 715 bp in length
* 47245 47344: gap of 100 bp
* 47345 48088: contig of 744 bp in length
* 48089 48188: gap of 100 bp
* 48189 48948: contig of 760 bp in length
* 48949 49048: gap of 100 bp
* 49049 49803: contig of 755 bp in length
* 49804 49903: gap of 100 bp
* 49904 50648: contig of 745 bp in length
* 50649 50748: gap of 100 bp
* 50749 51506: contig of 758 bp in length
* 51507 51606: gap of 100 bp
* 51607 52345: contig of 739 bp in length
* 52346 52445: gap of 100 bp
* 52446 53185: contig of 740 bp in length
* 53186 53286: gap of 100 bp
* 53286 54059: contig of 774 bp in length
* 54060 54159: gap of 100 bp
* 54160 54901: contig of 742 bp in length
* 54902 55001: gap of 100 bp
* 55002 55740: contig of 739 bp in length
* 55741 55840: gap of 100 bp
* 55841 56577: contig of 737 bp in length
* 56578 57433: contig of 756 bp in length
* 57434 57533: gap of 100 bp
* 57534 58290: contig of 757 bp in length

```

Query Match 17.3%; Score 22; DB 2; Length 59155;
 Best Local Similarity 100.0%; Pred. No. 0.058;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAAGAGACTGTGTGTGTCT 92
 |||

Db 7798 CAAGAGACTGTGTGTGTCT 7819

RESULT 50
 AC108040/c
 LOCUS Homo sapiens chromosome 4 clone RP11-210010, WORKING DRAFT
 DEFINITION
 AC108040
 AC108040
 AC108040.2 GI:18425316
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULFILL.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 163521)
 TITLE The sequence of Homo sapiens clone
 AUTHORS Unpublished
 JOURNAL 2 (bases 1 to 163521)
 REFERENCE Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (24-JAN-2002) Genome Sequencing Center, Washington
 TITLE University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 JOURNAL MO 63108, USA
 COMMENT On Jan 30, 2002 this sequence version replaced gi:18308830.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Contact: submissions@watson.wustl.edu
 ----- Project Information -----

Center project name: H_NH0210010
 ----- Summary Statistics -----
 Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-Primer ET; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 162304 bases at least Q40
 Consensus quality: 162444 bases at least Q30
 Consensus quality: 162529 bases at least Q20
 Insert size: 189000; agarose-fp
 Insert size: 163221; sum-of-contigs
 Quality coverage: 8.88 in Q20 bases; agarose-fp
 Quality coverage: 8.85 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1564: contig of 1564 bp in length
 * 1565 1664: gap of unknown length
 * 1665 88984: contig of 87320 bp in length
 * 88985 89084: gap of unknown length
 * 89085 112702: contig of 23618 bp in length
 * 112703 112803: gap of unknown length
 * 112803 163521: contig of 50719 bp in length.
 FEATURES
 Location/Qualifiers
 1. 163521
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="RP11-210010"
 1. 1564
 /note="assembly_name:Contig25"
 clone end:SP6
 vector side:left
 1665 88984
 /note="assembly_name:Contig28"
 clone end:T7
 vector side:left
 89085 112702
 /note="assembly_name:Contig26"
 112803 163521
 /note="assembly_name:Contig27"
 BASE COUNT 51049 a 32426 c 33346 g 46399 t 301 others
 ORIGIN

Query Match 17.3%; Score 22; DB 2; Length 163521;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAAGAGACTGTGTGTGTCT 92
 |||

Db 80211 CAAGAGACTGTGTGTGTCT 80190

Search completed: November 7, 2003, 13:29:05
 Job time : 847.956 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 02:39:56 ; Search time 129.53 Seconds
(without alignments)
2646.715 Million cell updates/sec

Title: US-09-509-779-1_COPY_154_280

Perfect score: 127
Sequence: 1 GTGCGATACCTGTGCTCATCT.....ATTCTTCACACACTGCTGC 127

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N.Geneseq_190N03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	100.0	1140	20	AAx87313
2	32	25.2	224	20	AAx41085
3	32	25.2	342	21	AAx6883
4	32	25.2	342	22	AAH43570
5	32	25.2	342	24	AAH3674
6	32	25.2	754	20	AAx87317
7	32	25.2	754	20	AAx87321
8	32	25.2	754	20	AAx87322

9	32	25.2	754	20	AAx87323	Human sensitive to
10	32	25.2	754	20	AAx87324	Human sensitive to
11	32	25.2	754	20	AAx87325	Human sensitive to
12	32	25.2	754	20	AAx87326	Human sensitive to
13	32	25.2	754	20	AAx87327	Human sensitive to
14	32	25.2	754	20	AAx87328	Human sensitive to
15	32	25.2	754	20	AAx87330	Human sensitive to
16	32	25.2	754	20	AAx87331	Human sensitive to
17	32	25.2	754	20	AAx87314	Human sensitive to
18	32	25.2	836	22	AAx25860	Human sensitive to
19	32	25.2	836	22	AAx26319	Human sensitive to
20	32	25.2	836	25	ABX73201	Human sensitive to
21	32	25.2	836	25	ABX73660	Human sensitive to
22	32	25.2	962	21	AAx7504	Human sensitive to
23	30	23.6	754	20	AAx87320	Human sensitive to
24	30	23.6	754	20	AAx87329	Human sensitive to
25	29	22.8	747	20	AAx87315	Human sensitive to
26	29	22.8	754	20	AAx87318	Human sensitive to
27	29	22.8	754	20	AAx87319	Human sensitive to
28	29	22.8	1152	22	AAx25847	Human sensitive to
29	29	22.8	60	24	ABN40537	Human sensitive to
30	25	19.7	439	24	ABZ11414	Human sensitive to
31	25	19.7	441	21	AAx7493	Human sensitive to
32	25	19.7	706	20	AAx87316	Human sensitive to
33	21	16.5	264	22	AAx16209	Human sensitive to
34	21	16.5	596	22	AAx25052	Human sensitive to
35	21	16.5	630	22	AAx22239	Human sensitive to
36	19	15.0	225	24	ABQ90662	Human sensitive to
37	19	15.0	498	22	AAx36982	Human sensitive to
38	19	15.0	498	25	ABX59970	Human sensitive to
39	18	14.2	18	20	AAx87337	Human sensitive to
40	17	13.4	239	23	AAx84630	Human sensitive to
41	17	13.4	348	25	ABx20540	Human sensitive to
42	17	13.4	355	22	ABA46329	Human sensitive to
43	17	13.4	355	22	ABA51431	Human sensitive to
44	17	13.4	355	22	ABA56889	Human sensitive to
45	17	13.4	355	22	ABA69454	Human sensitive to

ALIGNMENTS

RESULT 1

ID AAX87313 standard; cDNA, 1140 BP.

AC AAX87313;

DT 27-SEP-1999 (first entry)

XX Mouse sensitive to apoptosis (SAG) gene.

DE SAG gene; sensitive to apoptosis; mouse; cancer; tumour;

XX neurodegenerative disease; muscular dystrophy; wound healing;

KW vulnerability; therapy; ds.

XX Mus musculus.

OS Mus musculus.

XX Key Location/Qualifiers

FT CDS 17..358

FT /tag= a

XX W09932514-A2.

XX PD 01-JUL-1999.

XX PF 15-DEC-1998; 98WO-US26705.

XX PR 11-SEP-1998; 98US-00998490

XX PR 19-DEC-1997; 97US-00681179

XX (WARN) WARNER LAMBERT CO.

```

PI Sun Y;
XX
XX WPI: 1999-430152/36.
DR P-PSDB; AAY06491.
XX
XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
XX
XX Claim 1; Page 48-49; 84pp; English.
XX
XX This is the nucleotide sequence of mouse sensitive to apoptosis
CC (SAG) cDNA, which codes for a novel redox-sensitive, haem-binding
CC protein (see AAY06491) with a zinc RING finger domain that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversal of a tumour phenotype.
CC The cDNA was isolated using differential display to identify genes
CC associated with 1,10-phenanthroline (OP)-induced apoptosis in
CC murine tumour lines, and use of an isolated OP-inducible clone to
CC screen a mouse lung cDNA library. The mouse SAG cDNA was used to
CC identify human SAG (see AAX87314). SAG is highly conserved among
CC species. Disruption in yeast was shown to be lethal. SAG deletion
CC mutants (see AAX87315-16) have been identified in human cancer lines,
CC suggesting a role in carcinogenesis. SAG genes, and mutant SAG
CC genes, can be used to protect cells from apoptosis induced by redox
CC reagents. Antisense SAG genes can be used to inhibit the growth of
CC tumour cells. The SAG genes can also be used for the recombinant
CC production of the SAG proteins. The SAG proteins can be used to
CC scavenge oxygen radicals in organisms and to promote wound healing.
CC Additionally, the SAG genes or their complements can be used to
CC promote or inhibit the growth of plant cells (all claimed). The SAG
CC protein is also an ideal molecular target in the development of
CC drugs against neurodegenerative disorders, cancers and muscle
CC dystrophy.
XX
XX Sequence 1140 BP; 302 A; 224 C; 287 G; 327 T; 0 other;
SQ
Query Match 100.0%; Score 127; DB 20; Length 1140;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GTGCGATACCTGTGCGATCTGCAGGGTCCAGGTGATGCGCTTCGATGTCAGC 60
DB 154 GTGCGATACCTGTGCGATCTGCAGGGTCCAGGTGATGCGCTTCGATGTCAGC 213
OY 61 TGAANAACAACAAGAGACTGTGTGTGCTCTGGGGAAGTGTAACTTCCTTCACAA 120
DB 214 TGAANAACAACAAGAGACTGTGTGTGCTCTGGGGAAGTGTAACTTCCTTCACAA 273
OY 121 CTGCTGC 127
DB 274 CTGCTGC 280

```

```

XX
XX 11-FEB-1999.
PD
XX
XX 31-JUL-1998; 98MO-IB01238.
PF
XX
XX 01-AUG-1997; 97US-0905134.
PR
XX
XX (GENSET) GENSET.
PA
XX
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
PI
XX
XX WPI: 1999-153784/13.
DR
XX
XX P-PSDB; AAY12252.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from kidney, fetal kidney, dystrophic
PT muscle, muscle and heart tissue
XX
XX Claim 1; Page 441; 622pp; English.
PS
XX
XX AAY40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY01602 and
CC AAY11994 to AAY12260, respectively. The proteins given represent the
CC signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used
CC for directing extracellular secretion of a polypeptide or the insertion
CC of a polypeptide into a membrane, or importing a polypeptide into
CC a cell.
XX
XX Sequence 224 BP; 41 A; 60 C; 83 G; 38 T; 2 other;
SQ
Query Match 25.2%; Score 32; DB 20; Length 224;
Best Local Similarity 100.0%; Pred. No. 4.6e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45
DB 174 GCCATCTGCAGGGTCCAGGTGATGATGCTG 205

```

```

RESULT 2
AAX41085
ID AAX41085 standard; cDNA; 224 BP.
XX
XX AAX41085;
AC
XX
XX 18-JUN-1999 (first entry)
DT
XX
XX Human secreted protein 5' EST SEQ ID NO: 297.
DE
XX
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
XX Homo sapiens.
OS
XX
XX MO9906554-A2.
PN

```

```

RESULT 3
AAA96883
ID AAA96883 standard; DNA; 342 BP.
XX
XX AAA96883;
AC
XX
XX 19-FEB-2001 (first entry)
DT
XX
XX Nucleotide sequence of human ring finger protein ROC2.
DE
XX
XX ROC1; ROC2; cullin; ring finger protein; APC1; APC complex; SCF pathway;
KW cullin dependent ubiquitin ligase; CDK inhibitor Sic1 degradation;
KW tumour; ss.
XX
XX
XX Homo sapiens.
OS
XX
XX Key
FH CDS
FT 1..342
FT Location/Qualifiers
FT /*tag= a
FT /product= "ROC2"
FT /transl_except= (pos: 28..30, aa: Pro)
FT /transl_except= (pos: 34..36, aa: Val)
FT /transl_except= (pos: 40..42, aa: Ser)

```

```

FT      /cransl_except= (pos: 58..60, aa: Ala)
PT      /cransl_except= (pos: 67..69, aa: Lys)
FT      /cransl_except= (pos: 109..111, aa: Ala)
XX
XX      MO200058472-A2.
XX
XX      05-OCT-2000.
XX
XX      31-MAR-2000; 2000MO-US08592.
XX
XX      31-MAR-1999; 99US-0127261.
XX      22-NOV-1999; 99US-0166927.
XX
XX      (UNNC-) UNIV NORTH CAROLINA.
XX
XX      Xiong Y, Ohta T;
XX      WPI; 2000-647235/62.
XX      P-PSDB; AAB19161.
XX
XX      Novel nucleic acid encoding cullin regulating ring finger proteins,
XX      termed as ROC proteins similar to anaphase-promoting complex 11, for
XX      therapeutic and diagnostic use
XX      Claim 18; Fig 2B; 83pp; English.
XX
XX      The present sequence encodes a human ROC2 ring finger protein. The
XX      specification also describes human ROC1, ROC1 and ROC2 are similar
XX      to APC11, a subunit of the APC complex. The proteins stimulate cullin
XX      dependent ubiquitin ligase activity. ROC1 functions in vivo as an
XX      essential regulator of CDK inhibitor Sic1 degradation by the SCF
XX      (undefined) pathway. ROC proteins are useful for screening bioactive
XX      agents that interfere with the binding of ROC proteins with cullin
XX      proteins. Pharmaceutical formulations comprising ROC proteins are
XX      useful for diagnostic and therapeutic purposes, preferably for
XX      diagnosing and treating tumours.
XX
XX      Sequence 342 BP; 82 A; 84 C; 108 G; 68 T; 0 other;
XX
XX      Query Match      25.2%; Score 32; DB 21; Length 342;
XX      Best Local Similarity 100.0%; Pred. No. 4.5e-07;
XX      Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45
XX      |||
XX      151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182
XX
XX      RESULT 4
XX      ID      AAH43570
XX      ID      AAH43570 standard; cDNA; 342 BP.
XX
XX      AC      AAH43570;
XX
XX      DT      07-JAN-2002 (first entry)
XX
XX      DE      ROC2 coding sequence.
XX
XX      KW      Assay; ubiquitin ligase; tag1-ubiquitin; E1; E2;
XX      ubiquitin activating enzyme; ubiquitin conjugating enzyme; E3;
XX      ubiquitin ligase; ubiquitination modulator; ss.
XX
XX      OS      Unidentified.
XX
XX      PN      WO200175145-A2.
XX
XX      PD      11-OCT-2001.
XX
XX      PF      03-APR-2001; 2001WO-US10906.
XX      03-APR-2000; 2000US-0542497.
XX
XX      PA      (RIGB-) RIGEL PHARM INC.

```

```

XX      Issakani SD, Huang J, Sheung J, Pray TR;
XX      WPI; 2001-626445/72.
XX      P-PSDB; AAB47600.
XX
XX      Assaying ubiquitin ligase activity for identifying modulators of
XX      ubiquitination, by combining ubiquitin, ubiquitin activating
XX      conjugating enzyme, ubiquitin ligase and measuring amount of ubiquitin
XX      bound to the ligase
XX
XX      Example 1; Fig 12A; 98pp; English.
XX
XX      The sequences given in AAH43568-72 encode proteins which may be used
XX      in the method of the invention for assaying ubiquitin ligase activity.
XX      The method comprises combining under conditions that favour ubiquitin
XX      ligase activity, tag1-ubiquitin, E1 (ubiquitin activating enzyme), E2
XX      (ubiquitin conjugating enzyme) and E3 (ubiquitin ligase) and measuring
XX      the amount of tag1-ubiquitin bound to the E3. The method is useful for
XX      assaying ubiquitin ligase activity and ubiquitination enzyme activity
XX      which is useful for identifying ubiquitination modulator. The method
XX      comprises combining tag1-ubiquitin, the modulator, E1, E2 and tag2-E3
XX      and measuring the amount of tag1-ubiquitin bound to tag2-E3 or combining
XX      tag1-ubiquitin, modulator, E1 and tag3-E2 and measuring the amount of
XX      tag1-ubiquitin bound to tag3-E2. Ubiquitin ligase activity is measured
XX      directly, eliminating the need for target proteins and subsequent
XX      analysis such as separating ligated from unligated material in an
XX      SDS-PAGE procedure. This allows multi-well array analysis and high
XX      throughput screening techniques for modulators of ubiquitination
XX      activity. The method also allows the analysis of many different
XX      combinations of E3 components and E2/E3 combinations without requiring
XX      prior identification of specific target substrates. Ubiquitin is labeled,
XX      directly or indirectly and this allows for easy and rapid detection and
XX      measurement of ligated ubiquitin.
XX
XX      Sequence 342 BP; 82 A; 84 C; 108 G; 68 T; 0 other;
XX
XX      Query Match      25.2%; Score 32; DB 22; Length 342;
XX      Best Local Similarity 100.0%; Pred. No. 4.5e-07;
XX      Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45
XX      |||
XX      151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182
XX
XX      RESULT 5
XX      ID      AAD39674
XX      ID      AAD39674 standard; DNA; 342 BP.
XX
XX      AC      AAD39674;
XX
XX      DT      22-OCT-2002 (first entry)
XX
XX      DE      Human RING finger protein, ROC2 DNA.
XX
XX      KW      Ubiquitin ligase; U1; tag1-ubiquitin; E1; ubiquitin-activating enzyme;
XX      E2; ubiquitin-conjugating enzyme/ubiquitin carrier protein; E3; human;
XX      RING finger protein; gene; ds.
XX
XX      OS      Homo sapiens.
XX
XX      PH      Key      Location/Qualifiers
XX      FT      CDS      1..342
XX      FT      /*tag= a
XX      FT      /product= "Human ROC2 protein"
XX
XX      PN      US2002042083-A1.
XX
XX      PD      11-APR-2002.
XX
XX      PF      03-APR-2001; 2001US-0826312.
XX
XX

```



```

PR 03-APR-2000; 2000US-0542497.
XX
XX (RICE-) RIGEL PHARM INC.
XX
XX Iseakani SD, Huang J, Sheung J, Pray TR;
XX WPI, 2002-488718/52.
XX DR P-PSDB; AAE24621.
XX
XX Assay for ubiquitin ligase activity, useful for identifying modulators,
XX by measuring binding of labeled ubiquitin to ubiquitin ligase
XX
XX Disclosure; Fig 12A; 56pp; English.
XX
XX The invention relates to an assay for ubiquitin ligase (UL) activity
XX which comprises incubating tag1-ubiquitin, E1 (ubiquitin-activating
XX enzyme), E2 (ubiquitin-conjugating enzyme/ubiquitin carrier protein)
XX and E3 (UL) and measuring the amount of tag1-ubiquitin bound to E3. The
XX method is particularly used to screen for modulators of UL activity. The
XX present sequence is human RING finger protein, ROC2 DNA.
XX
XX Sequence 342 BP; 82 A; 84 C; 108 G; 68 T; 0 other;
XX
XX Query Match 25.2%; Score 32; DB 24; Length 342;
XX Best Local Similarity 100.0%; Pred. No. 4.5e-07;
XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 14 GCCATCTGCAGGCTCCAGGTGATGATGCCTG 45
XX 151 GCCATCTGCAGGCTCCAGGTGATGATGCCTG 182
XX
XX RESULT 6
XX ID AAX87317 standard; cDNA; 754 BP.
XX
XX AAX87317;
XX
XX 27-SEP-1999 (first entry)
XX
XX Human sensitive to apoptosis (SAG) gene mutant MML.
XX
XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnerability; therapy; mutant; ds.
XX
XX Homo sapiens.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 1..342
XX FT /*tag= a
XX FT replace(148,T)
XX FT /*tag= b
XX FT /note= "C50S mutation"
XX
XX WO9932514-A2.
XX
XX 01-JUL-1999.
XX
XX 15-DEC-1998; 98WO-US26705.
XX
XX 11-SEP-1998; 98US-0099840.
XX 19-DEC-1997; 97US-0068179.
XX
XX (WARN ) WARNER LAMBERT CO.
XX
XX Sun Y;
XX
XX WPI; 1999-430152/36.
XX DR P-PSDB; AAY06495.
XX
XX SAG; Sensitive to Apoptosis Gene and related proteins, useful for

```

```

PT promoting cell growth and protecting cells against apoptosis
XX
XX Claim 15; Page 59; 84pp; English.
XX
XX This is the nucleotide sequence of human sensitive to apoptosis
XX (SAG) mutant gene MML, which codes for a SAG protein (see AAY06495)
XX in which the Cys residue at position 50 of the native protein (see
XX AAY06492) is replaced by a Ser residue obtained by site-directed mutagenesis of SAG
XX 50 from TGC to AGC obtained by site-directed mutagenesis of SAG.
XX cDNA. This residue is at a haem binding site of SAG. Single and
XX double SAG mutants (see AAX87317-31) were made in order to determine
XX the role of each cysteine residue of SAG in haem binding and SAG
XX oligomerization. MML showed reduced haem binding but unaffected
XX cell growth, protects cells from apoptosis, scavenges oxygen
XX radicals and can be used for the reversion of a tumour phenotype.
XX SAG gene, and mutant SAG genes, can be used to protect cells from
XX apoptosis induced by redox reagents. They can also be used for the
XX recombinant production of SAG proteins, which are molecular targets
XX in the development of drugs against neurodegenerative disorders,
XX cancers and muscle dystrophy, and promoting wound healing.
XX
XX Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
XX
XX Query Match 25.2%; Score 32; DB 20; Length 754;
XX Best Local Similarity 100.0%; Pred. No. 4.5e-07;
XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 14 GCCATCTGCAGGCTCCAGGTGATGATGCCTG 45
XX 151 GCCATCTGCAGGCTCCAGGTGATGATGCCTG 182
XX
XX RESULT 7
XX ID AAX87321 standard; cDNA; 754 BP.
XX
XX AAX87321;
XX
XX 27-SEP-1999 (first entry)
XX
XX Human sensitive to apoptosis (SAG) gene mutant MMS.
XX
XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnerability; therapy; mutant; ds.
XX
XX Homo sapiens.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 1..342
XX FT /*tag= a
XX FT replace(238,T)
XX FT /*tag= b
XX FT /note= "C80S mutation"
XX
XX WO9932514-A2.
XX
XX 01-JUL-1999.
XX
XX 15-DEC-1998; 98WO-US26705.
XX
XX 11-SEP-1998; 98US-0099840.
XX 19-DEC-1997; 97US-0068179.
XX
XX (WARN ) WARNER LAMBERT CO.
XX
XX Sun Y;
XX
XX WPI; 1999-430152/36.
XX DR P-PSDB; AAY06499.
XX
XX

```

PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
XX
XX
PS Claim 15; Page 66; 84pp; English.

CC This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM5, which codes for a SAG protein (see AAY06499)
CC in which the Cys residue at position 80 of the native protein (see
CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
CC 80 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is in zinc-binding finger 1 of SAG. Single and
CC double SAG mutants (see AAX87317-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM5
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversal of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.

SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 25.2%; Score 32; DB 20; Length 754;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45
DB 151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182

RESULT 8
AAX87322
ID AAX87322 standard; cDNA; 754 BP.
XX
XX AAX87322;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene mutant MM6.
XX
XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnerability; therapy; mutant; ds.
XX
XX Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..342
FT /*tag= a
FT replace(244..246,CAT)
FT /*tag= b
FT /note= "H82K mutation"
FT
FT
PN WO9932514-A2.
XX
XX 01-JUL-1999.
XX
XX
XX 15-DEC-1998; 98WO-US26705.
XX
XX 11-SEP-1998; 98US-0099840.
XX 19-DEC-1997; 97US-0068179.
XX
XX (WARN) WARNER LAMBERT CO.
XX
XX Sun Y;
XX
XX WPI; 1999-430152/36.
XX
XX P-PsDB; AAY06500.
DR

XX
XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
PT
XX
XX
PS Claim 15; Page 67-68; 84pp; English.

CC This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM6, which codes for a SAG protein (see AAY06500)
CC in which the His residue at position 82 of the native protein (see
CC AAY06492) is replaced by a Lys residue owing to a mutation of codon
CC 82 from CAT to AAA obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is in zinc-binding finger 1 of SAG. Single and
CC double SAG mutants (see AAX87317-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM6
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversal of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.

SQ Sequence 754 BP; 207 A; 154 C; 201 G; 192 T; 0 other;

Query Match 25.2%; Score 32; DB 20; Length 754;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45
DB 151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182

RESULT 9
AAX87323
ID AAX87323 standard; cDNA; 754 BP.
XX
XX AAX87323;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene mutant MM7.
XX
XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnerability; therapy; mutant; ds.
XX
XX Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..342
FT /*tag= a
FT replace(253..255,CAC)
FT /*tag= b
FT /note= "H85K mutation"
FT
FT
PN WO9932514-A2.
XX
XX 01-JUL-1999.
XX
XX
XX 15-DEC-1998; 98WO-US26705.
XX
XX 11-SEP-1998; 98US-0099840.
XX 19-DEC-1997; 97US-0068179.
XX
XX (WARN) WARNER LAMBERT CO.
XX
XX Sun Y;
XX
XX WPI; 1999-430152/36.
XX
XX P-PsDB; AAY06500.
DR

DR P-PSDB; AAY06501.

XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
XX
XX
PS Claim 15; Page 69-70; 84pp; English.

CC This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM7, which codes for a SAG protein (see AAY06501)
CC in which the His residue at position 85 of the native protein (see
CC AAY06492) is replaced by a Lys residue owing to a mutation of codon
CC 85 from CAC to CAG obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is in zinc-binding finger 2 of SAG. Single and
CC double SAG mutants (see AAX87317-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM7
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversal of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.

XX Sequence 754 BP; 206 A; 153 C; 202 G; 193 T; 0 other;

Query Match 25.2%; Score 32; DB 20; Length 754;

Best Local Similarity 100.0%; Pred. No. 4.5e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45

DB 151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182

RESULT 10

AAX87324

ID AAX87324 standard; cDNA; 754 BP.

XX AAX87324;

DT 27-SEP-1999 (first entry)

XX Human sensitive to apoptosis (SAG) gene mutant MM8.

XX SAG gene; sensitive to apoptosis; human; cancer; tumour;

KW neurodegenerative disease; muscular dystrophy; wound healing;

KW vulnerability; therapy; mutant; ds.

XX Homo sapiens.

OS Synthetic.

XX Key

FT CDS 1..342 Location/Qualifiers

FT mutation /tag= a

FT /replace(262,T)

FT /tag= b

FT /note= "C88S mutation"

XX MO9932514-A2.

XX 01-JUL-1999.

XX 15-DEC-1998; 98WO-US26705.

XX 11-SEP-1998; 98US-0099840.

XX 19-DEC-1997; 97US-0068179.

XX (WARN) WARNER LAMBERT CO.

XX PA Sun Y;

XX PI

XX

DR WPI; 1999-430152/36.

XX P-PSDB; AAY06502.

XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
XX
XX
PS Claim 15; Page 71; 84pp; English.

CC This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM8, which codes for a SAG protein (see AAY06502)
CC in which the Cys residue at position 88 of the native protein (see
CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
CC 88 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is in zinc-binding finger 2 of SAG. Single and
CC double SAG mutants (see AAX87317-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM8
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversal of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.

XX Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 25.2%; Score 32; DB 20; Length 754;

Best Local Similarity 100.0%; Pred. No. 4.5e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGGTCCAGGTGATGATGCTG 45

DB 151 GCCATCTGCAGGGTCCAGGTGATGATGCTG 182

RESULT 11

AAX87325

ID AAX87325 standard; cDNA; 754 BP.

XX AAX87325;

DT 27-SEP-1999 (first entry)

XX Human sensitive to apoptosis (SAG) gene mutant MM9.

XX SAG gene; sensitive to apoptosis; human; cancer; tumour;

KW neurodegenerative disease; muscular dystrophy; wound healing;

KW vulnerability; therapy; mutant; ds.

XX Homo sapiens.

OS Synthetic.

XX Key

FT CDS 1..342 Location/Qualifiers

FT mutation /tag= a

FT /replace(295,T)

FT /tag= b

FT /note= "C99S mutation"

XX MO9932514-A2.

XX 01-JUL-1999.

XX 15-DEC-1998; 98WO-US26705.

XX 11-SEP-1998; 98US-0099840.

XX 19-DEC-1997; 97US-0068179.

XX (WARN) WARNER LAMBERT CO.

XX PA Sun Y;

XX PI

XX

XX MPI; 1999-430152/36.
DR P-PSDB; AAY06503.
XX
PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PS promoting cell growth and protecting cells against apoptosis
XX Claim 15; Page 72-73; 84pp; English.
XX This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM9, which codes for a SAG protein (see AAY06503)
CC in which the Cys residue at position 99 of the native protein (see
CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
CC 99 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is in zinc-finger finger 2 of SAG. Single and
CC double SAG mutants (see AAX87317-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM9
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversion of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.
XX
SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
Query Match 25.2%; Score 32; DB 20; Length 754;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 GCCATCTGCAGGCTCCAGTGATGATGCCTG 45
DB 151 GCCATCTGCAGGCTCCAGTGATGATGCCTG 182
RESULT 12
AAX87326
ID AAX87326 standard; cDNA; 754 BP.
XX
AC AAX87326;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene mutant MM10.
XX
KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
KW neurodegenerative disease; muscular dystrophy; wound healing;
KW vulnerability; therapy; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..342
FT mutation /tag= a
FT mutation /tag= b
FT mutation /note= "H85K mutation"
FT mutation replace(253..255,CAC)
FT /tag= C
FT /note= "H85K mutation"
XX
XX WO9932514-A2.
XX
XX 01-JUL-1999.
XX
XX 15-DEC-1998; 98WO-US26705.
XX
XX 11-SEP-1998; 98US-0099840.
XX
PR 19-DEC-1997; 97US-0068179.

XX (WARN) WARNER LAMBERT CO.
PA
XX Sun Y;
XX
XX MPI; 1999-430152/36.
DR P-PSDB; AAY06504.
XX
PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PS promoting cell growth and protecting cells against apoptosis
XX Claim 15; Page 74-75; 84pp; English.
XX This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM10, which codes for a SAG protein (see AAY06504)
CC in which the His residues at positions 82 and 85 of the native
CC protein (see AAY06492) are replaced by Lys residues owing to
CC mutations of codons 82 and 85 by site-directed mutagenesis of SAG
CC cDNA. These residues are in zinc-finger fingers 1 and 2 of SAG. SAG
CC mutants (see AAX87317-31) were made in order to determine the role of
CC each cysteine residue of SAG in haem binding and oligomerization.
CC These properties were unaffected by the MM10 mutations. SAG is a
CC novel zinc finger protein that promotes cell growth, protects cells
CC from apoptosis, scavenges oxygen radicals and can be used for the
CC reversion of a tumour phenotype. SAG genes, and mutant SAG genes,
CC can be used to protect cells from apoptosis induced by redox
CC reagents. They can also be used for the recombinant production of
CC SAG proteins, which are molecular targets in the development of
CC drugs against neurodegenerative disorders, cancers and muscle
CC dystrophy, and promoting wound healing.
XX
SQ Sequence 754 BP; 208 A; 152 C; 202 G; 192 T; 0 other;
Query Match 25.2%; Score 32; DB 20; Length 754;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 GCCATCTGCAGGCTCCAGTGATGATGCCTG 45
DB 151 GCCATCTGCAGGCTCCAGTGATGATGCCTG 182
RESULT 13
AAX87327
ID AAX87327 standard; cDNA; 754 BP.
XX
AC AAX87327;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene mutant MM11.
XX
KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
KW neurodegenerative disease; muscular dystrophy; wound healing;
KW vulnerability; therapy; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..342
FT mutation /tag= a
FT mutation /tag= b
FT mutation replace(190,T)
FT /tag= C
FT /note= "C64S mutation"
XX
XX WO9932514-A2.
XX
XX 01-JUL-1999.
XX
XX 15-DEC-1998; 98WO-US26705.
XX
XX 11-SEP-1998; 98US-0099840.

```

PR 19-DEC-1997; 97US-0068179.
XX
XX (WARN ) WARNER LAMBERT CO.
XX
XX Sun Y;
XX
XX WPI; 1999-430152/36.
XX
XX P-PSDB; AAY06505.
XX
XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
XX promoting cell growth and protecting cells against apoptosis
XX
XX Claim 15; Page 76; 84pp; English.
XX
XX This is the nucleotide sequence of human sensitive to apoptosis
XX (SAG) mutant gene MM11, which codes for a SAG protein (see AAY06505)
XX in which the Cys residue at position 64 of the native protein (see
XX AAY06492) is replaced by a Ser residue owing to a mutation of codon
XX 64 from TGT to AGC obtained by site-directed mutagenesis of SAG
XX cDNA. This residue is in zinc-finger 1 of SAG. Single and
XX double SAG mutants (see AAX87317-31) were made in order to determine
XX the role of each cysteine residue of SAG in haem binding and SAG
XX oligomerization. These properties were unaffected by the MM11
XX mutation. SAG is a novel zinc finger protein that promotes
XX cell growth, protects cells from apoptosis, scavenges oxygen
XX radicals and can be used for the reversion of a tumour phenotype.
XX SAG genes, and mutant SAG genes, can be used to protect cells from
XX apoptosis induced by redox reagents. They can also be used for the
XX recombinant production of SAG proteins, which are molecular targets
XX in the development of drugs against neurodegenerative disorders,
XX cancers and muscle dystrophy, and promoting wound healing.
XX
XX Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
XX
XX Query Match 25.2%; Score 32; DB 20; Length 754;
XX Best Local Similarity 100.0%; Pred. No. 4.5e-07;
XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 14 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 45
DB 151 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 182

```

```

PR 11-SEP-1998; 98US-0099840.
XX
XX 19-DEC-1997; 97US-0068179.
XX
XX (WARN ) WARNER LAMBERT CO.
XX
XX Sun Y;
XX
XX WPI; 1999-430152/36.
XX
XX P-PSDB; AAY06506.
XX
XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
XX promoting cell growth and protecting cells against apoptosis
XX
XX Claim 15; Page 77-78; 84pp; English.
XX
XX This is the nucleotide sequence of human sensitive to apoptosis
XX (SAG) mutant gene MM12, which codes for a SAG protein (see AAY06506)
XX in which the Cys residue at position 73 of the native protein (see
XX AAY06492) is replaced by a Ser residue owing to a mutation of codon
XX 73 from TGT to AGC obtained by site-directed mutagenesis of SAG
XX cDNA. This residue is in a proase inhibitor motif of SAG. Single
XX and double SAG mutants (see AAX87317-31) were made to determine
XX the role of each cysteine residue of SAG in haem binding and SAG
XX oligomerization. These properties were unaffected by the MM12
XX mutation. SAG is a novel zinc finger protein that promotes
XX cell growth, protects cells from apoptosis, scavenges oxygen
XX radicals and can be used for the reversion of a tumour phenotype.
XX SAG genes, and mutant SAG genes, can be used to protect cells from
XX apoptosis induced by redox reagents. They can also be used for the
XX recombinant production of SAG proteins, which are molecular targets
XX in the development of drugs against neurodegenerative disorders,
XX cancers and muscle dystrophy, and promoting wound healing.
XX
XX Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;
XX
XX Query Match 25.2%; Score 32; DB 20; Length 754;
XX Best Local Similarity 100.0%; Pred. No. 4.5e-07;
XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 14 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 45
DB 151 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 182

```

```

RESULT 14
AAX87328
ID AAX87328 standard; cDNA; 754 BP.
XX
XX AAX87328;
XX
XX 27-SEP-1999 (first entry)
XX
XX Human sensitive to apoptosis (SAG) gene mutant MM12.
XX
XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnery; therapy; mutant; ds.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FH 1..342
XX FT /*tag= a
XX FT replace(217,T)
XX FT /*tag= b
XX FT /note= "C73S mutation"
XX
XX WO9932514-A2.
XX
XX 01-JUL-1999.
XX
XX 15-DEC-1998; 98WO-US26705.
XX

```

```

RESULT 15
AAX87330
ID AAX87330 standard; cDNA; 754 BP.
XX
XX AAX87330;
XX
XX 27-SEP-1999 (first entry)
XX
XX Human sensitive to apoptosis (SAG) gene mutant MM14.
XX
XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnery; therapy; mutant; ds.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FH 1..342
XX FT /*tag= a
XX FT replace(295,C)
XX FT /*tag= b
XX FT /note= "C99S mutation"
XX FT replace(304,C)
XX FT /*tag= c
XX FT /note= "C102S mutation"
XX
XX WO9932514-A2.
XX

```

PD 01-JUL-1999.
XX
PF 15-DEC-1998; 98WO-US26705.
XX
PR 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.
XX
PA (WARN) WARNER LAMBERT CO.
XX
PI Sun Y;
XX WPI; 1999-430152/36.
DR P-PSDB; AAY06508.
XX
PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
XX
PS Claim 15; Page 81; 84pp; English.
XX
XX This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM14, which codes for a SAG protein (see AAY06508)
CC in which the Cys residues at positions 99 and 102 of the native
CC protein (see AAY06492) are replaced by Ser residues owing to
CC mutations of codons 99 and 102 through site-directed mutagenesis of
CC SAG cDNA. These residues are in zinc ring finger 2 of SAG. Single
CC double SAG mutants (see AAX87317-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. Haem binding was unaffected in MM14, but no
CC oligomerization was observed, indicating that these residues are
CC important for intermolecular disulfide bond formation. SAG is
CC a novel zinc finger protein that promotes cell growth, protects cells
CC from apoptosis, scavenges oxygen radicals and can be used for the
CC reversion of a tumour phenotype. SAG genes, and mutant SAG genes,
CC can be used to protect cells from apoptosis induced by redox
CC reagents. They can also be used for the recombinant production of
CC SAG proteins, which are molecular targets in the development of
CC drugs against neurodegenerative disorders, cancers and muscle
CC dystrophy, and promoting wound healing.
XX
SO Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;

Query Match 25.2%; Score 32; DB 20; Length 754;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGTCAGGTGATGATGCCTG 45
|||
DB 151 GCCATCTGCAGGTCAGGTGATGATGCCTG 182

RESULT 16
AAX87331
ID AAX87331 standard; cDNA; 754 BP.
XX
AC AAX87331;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene mutant MM15.
XX
KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
KW neurodegenerative disease; muscular dystrophy; wound healing;
KW vulnerability; therapy; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..342
FT mutation /*tag= a
FT /tag= b
FT /note= "C47S mutation"

XX
PN WO9932514-A2.
XX
PD 01-JUL-1999.
XX
PF 15-DEC-1998; 98WO-US26705.
XX
PR 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.
XX
PA (WARN) WARNER LAMBERT CO.
XX
PI Sun Y;
XX WPI; 1999-430152/36.
DR P-PSDB; AAY06509.
XX
PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
XX
PS Claim 15; Page 82-83; 84pp; English.
XX
XX This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM15, which codes for a SAG protein (see AAY06509)
CC in which the Cys residue at position 47 of the native protein (see
CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
CC 47 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is in a GADPH binding site of SAG. Single
CC and double SAG mutants (see AAX87317-31) were made to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM15
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversion of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.
XX
SO Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 25.2%; Score 32; DB 20; Length 754;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGTCAGGTGATGATGCCTG 45
|||
DB 151 GCCATCTGCAGGTCAGGTGATGATGCCTG 182

RESULT 17
AAX87314
ID AAX87314 standard; cDNA; 754 BP.
XX
AC AAX87314;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene.
XX
KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
KW neurodegenerative disease; muscular dystrophy; wound healing;
KW vulnerability; therapy; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..342
FT mutation /*tag= a
FT /tag= a
XX WO9932514-A2.

PD 01-JUL-1999.
 XX 15-DEC-1998; 98WO-US26705.
 PF 11-SEP-1998; 98US-0099840.
 XX 19-DEC-1997; 97US-0068179.
 PR (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 DR MPI: 1999-430152/36.
 XX P-PSDB; AAY06492.
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 XX promoting cell growth and protecting cells against apoptosis
 PS Claim 9; Page 50-51; 84pp; English.
 XX
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) cDNA, which codes for a novel redox-sensitive, haem-binding
 CC protein (see AAY06492) with a zinc RING finger domain that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC The cDNA was isolated from a HeLa cell cDNA library using mouse
 CC SAG cDNA (see AAX87313) as probe. SAG is highly conserved among
 CC species. Disruption in yeast was shown to be lethal. SAG deletion
 CC mutants (see AAX87315-16) have been identified in human cancer lines,
 CC suggesting a role in carcinogenesis. SAG genes, and mutant SAG
 CC genes, can be used to protect cells from apoptosis induced by redox
 CC reagents. Antisense SAG genes can also be used to inhibit the growth of
 CC tumour cells. The SAG proteins can also be used for the recombinant
 CC production of the SAG proteins. The SAG proteins can be used to
 CC scavenge oxygen radicals in organisms and to promote wound healing.
 CC Additionally, the SAG genes or their complements can be used to
 CC promote or inhibit the growth of plant cells (all claimed). The SAG
 CC protein is also an ideal molecular target in the development of
 CC drugs against neurodegenerative disorders, cancers and muscle
 CC dystrophy.
 XX
 SQ Sequence 754 BP; 205 A; 155 C; 201 G; 193 T; 0 other;
 Query Match 25.2%; Score 32; DB 20; Length 754;
 Best local Similarity 100.0%; Pred.No. 4.5e-07;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 14 GCCATCGCAGGGTCGAGTGATGATGCTG 45
 DB 151 GCCATCGCAGGGTCGAGTGATGATGCTG 182
 RESULT 18
 AAS25860
 ID AAS25860 standard; cDNA; 836 BP.
 XX AAS25860;
 AC
 AC
 DT 07-NOV-2001 (first entry)
 XX
 DE Human cDNA encoding a novel secreted protein, Seq ID 39.
 XX
 XX Human; immunosuppressive; antiarthritis; ss; antirheumatic;
 KW cytoskeletal; cardiant; vasotropic; cerebroprotective; neotropic;
 KW neuroprotective; antibacterial; virocidic; fungicide; ophthalmological;
 KW vulnereary; secreted protein; rheumatoid arthritis;
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
 KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin ageing; food additive; preservative; antiproliferative.
 XX
 OS Homo sapiens.
 XX

PN WO200155322-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01341.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0231415.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0232402.
 PR 14-SEP-2000; 2000US-0232403.
 PR 14-SEP-2000; 2000US-0232404.
 PR 14-SEP-2000; 2000US-0232405.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 21-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.

PR	29-SEP-2000	2000US-02363627
PR	29-SEP-2000	2000US-02363637
PR	29-SEP-2000	2000US-02363638
PR	29-SEP-2000	2000US-02363639
PR	29-SEP-2000	2000US-02363640
PR	02-OCT-2000	2000US-02363682
PR	02-OCT-2000	2000US-02370337
PR	02-OCT-2000	2000US-02370338
PR	02-OCT-2000	2000US-02370339
PR	02-OCT-2000	2000US-02370340
PR	13-OCT-2000	2000US-02393935
PR	13-OCT-2000	2000US-02393937
PR	20-OCT-2000	2000US-02409660
PR	20-OCT-2000	2000US-02412121
PR	20-OCT-2000	2000US-02411855
PR	20-OCT-2000	2000US-02411876
PR	20-OCT-2000	2000US-02411877
PR	20-OCT-2000	2000US-02418089
PR	20-OCT-2000	2000US-02441826
PR	01-NOV-2000	2000US-02446177
PR	08-NOV-2000	2000US-02446174
PR	08-NOV-2000	2000US-02446175
PR	08-NOV-2000	2000US-02446176
PR	08-NOV-2000	2000US-02446177
PR	08-NOV-2000	2000US-02446178
PR	08-NOV-2000	2000US-02446532
PR	08-NOV-2000	2000US-02446532
PR	08-NOV-2000	2000US-02446539
PR	08-NOV-2000	2000US-02446540
PR	08-NOV-2000	2000US-02446524
PR	08-NOV-2000	2000US-02446525
PR	08-NOV-2000	2000US-02446526
PR	08-NOV-2000	2000US-02446527
PR	17-NOV-2000	2000US-02496207
PR	17-NOV-2000	2000US-02496208
PR	17-NOV-2000	2000US-02496209
PR	17-NOV-2000	2000US-02496210
PR	17-NOV-2000	2000US-02496211
PR	17-NOV-2000	2000US-02496212
PR	17-NOV-2000	2000US-02496213
PR	17-NOV-2000	2000US-02496214
PR	17-NOV-2000	2000US-02496215
PR	17-NOV-2000	2000US-02496216
PR	17-NOV-2000	2000US-02496217
PR	17-NOV-2000	2000US-02496218
PR	17-NOV-2000	2000US-02496219
PR	17-NOV-2000	2000US-02496220
PR	17-NOV-2000	2000US-02496221
PR	17-NOV-2000	2000US-02496222
PR	17-NOV-2000	2000US-02496223
PR	17-NOV-2000	2000US-02496224
PR	17-NOV-2000	2000US-02496225
PR	17-NOV-2000	2000US-02496226
PR	17-NOV-2000	2000US-02496227
PR	17-NOV-2000	2000US-02496228
PR	17-NOV-2000	2000US-02496229
PR	17-NOV-2000	2000US-02496230
PR	17-NOV-2000	2000US-02496231
PR	17-NOV-2000	2000US-02496232
PR	17-NOV-2000	2000US-02496233
PR	17-NOV-2000	2000US-02496234
PR	17-NOV-2000	2000US-02496235
PR	17-NOV-2000	2000US-02496236
PR	17-NOV-2000	2000US-02496237
PR	17-NOV-2000	2000US-02496238
PR	17-NOV-2000	2000US-02496239
PR	17-NOV-2000	2000US-02496240
PR	17-NOV-2000	2000US-02496241
PR	17-NOV-2000	2000US-02496242
PR	17-NOV-2000	2000US-02496243
PR	17-NOV-2000	2000US-02496244
PR	17-NOV-2000	2000US-02496245
PR	17-NOV-2000	2000US-02496246
PR	17-NOV-2000	2000US-02496247
PR	17-NOV-2000	2000US-02496248
PR	17-NOV-2000	2000US-02496249
PR	17-NOV-2000	2000US-02496250
PR	17-NOV-2000	2000US-02496251
PR	17-NOV-2000	2000US-02496252
PR	17-NOV-2000	2000US-02496253
PR	17-NOV-2000	2000US-02496254
PR	17-NOV-2000	2000US-02496255
PR	17-NOV-2000	2000US-02496256
PR	17-NOV-2000	2000US-02496257
PR	17-NOV-2000	2000US-02496258
PR	17-NOV-2000	2000US-02496259
PR	17-NOV-2000	2000US-02496260
PR	17-NOV-2000	2000US-02496261
PR	17-NOV-2000	2000US-02496262
PR	17-NOV-2000	2000US-02496263
PR	17-NOV-2000	2000US-02496264
PR	17-NOV-2000	2000US-02496265
PR	17-NOV-2000	2000US-02496266
PR	17-NOV-2000</	

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM,

XX WPI; 2001-488783/53
DR P-PSDB; AAU15873.
DR
XX

PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -

Claim 1; SEQ ID No 39; 980pp; English

CC The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. Rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence encodes a novel secreted protein of the invention.

Query Match	25.2%	Score 32;	DB 22;	Length 836;
Best Local Similarity	100.0%	Pred. No. 4.5e-07;		
Matches 32; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY      14 GCCATCTGCAGGCTCCAGGTGATGCATGCCCTG 45
          |||||
Db      166 GCCATCTGCAGGCTCCAGGTGATGCATGCCCTG 197

```

RESULT 19

ID AAS26319 standard; cDNA; 836 BP

AC AAS26319;

DT 07-NOV-2001 (first entry)

DE Human cDNA encoding a novel secreted protein, Seq ID 498.

KM Human, immunosuppressive, antiarrhythmic, ss, anti-thematic;
 KM cytostatic; cardiac; vasotropic; cerebroprotective; nootropic;
 KM neuroprotective; antibacterial, viruside, fungicide; ophthalmological;
 KM vulnary; secreted protein; rheumatoid arthritis;
 KM hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KM cerebrovascular disorder; cerebral ischaemia; angio genesis;
 KM nervous system disorder; Alzheimer's disease; infection; ocular disorder;
 KM corneal infection; wound healing; epithelial cell proliferation;
 KM skin ageing; food additive; preservative; antiproliferative.

OS Homo sapiens.

PN WO200155322-A2

02-AUG-2001 PD

PF 17-JAN-2001; 2001WO-US01341.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234222.
PR 21-SEP-2000; 2000US-0234227.
PR 21-SEP-2000; 2000US-0234271.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239395.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.
PR 17-NOV-2000; 2000US-0249255.
PR 17-NOV-2000; 2000US-0249257.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251866.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251988.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-488783/53.
DR P-PSDB; AAUI6332.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -

XX Claim 1; SEQ ID No 498; 980bp; English.
PS
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention.

Query Match 25.2%; Score 32; DB 22; Length 836;
Best Local Similarity 100.0%; Pred. No. 4,5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

14 GCCATCTGCAGGCTCCAGTGTATGATGCTG 45
Db 219 GCCATCTGCAGGCTCCAGTGTATGATGCTG 250

RESULT 20
ABX73201
ID ABX73201 standard; DNA; 836 BP.
XX
AC ABX73201;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polynucleotide #29.
XX
KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotoxic; cytotoxic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
XX Homo sapiens.
XX
XX US2002132753-A1.
XX
PD 19-SEP-2002.
XX
PF 17-JAN-2001; 2001US-0764864.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217486P.
PR 14-JUL-2000; 2000US-218280P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.

PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225758P.
PR 14-AUG-2000; 2000US-225759P.
PR 22-AUG-2000; 2000US-226688P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.

(ROSE/) ROSEN C A.
(RUBE/) RUBEN S M.
(BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
XX P-PSDB; ABUS4941.
XX
XX New polypeptides and nucleic acids, useful in gene therapy for
XX treating, inhibiting or preventing e.g. neural, immune system,
XX muscular, respiratory, reproductive, gastrointestinal, pulmonary,
XX cardiovascular or renal disorders -
XX
XX
XX
PS Claim 1; SEQ ID NO 39; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
XX therapy for treating, inhibiting or preventing neural disorders, immune
XX system disorders (e.g. systemic lupus erythematosus, rheumatoid
XX arthritis and multiple sclerosis), muscular disorders, respiratory
XX diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis),
XX reproductive disorders, gastrointestinal disorders, pulmonary disorders,
XX cardiovascular disorders (e.g. congenital heart defects, Ebstein's
XX anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute
XX kidney failure and end-stage renal disease), hyperproliferative disorders
XX (e.g. Hodgkin's disease and leukaemia), inflammatory diseases (e.g.
XX septic shock, bursitis and appendicitis), allergic reactions and
XX conditions (e.g. asthma), blood related disorders (e.g. thrombosis,
XX atherosclerosis and myocardial infarction) and cancerous diseases.
XX Sequences ABX73173-ABX74167 represent human novel polynucleotides of the
XX invention.

```

XX SQ Sequence 836 BP; 264 A; 168 C; 209 G; 195 T; 0 other;
Query Match 25.2%; Score 32; DB 25; Length 836;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 14 GCCATCTGCAGGTCACAGTGTGATGATGCTTG 45
Db 166 GCCATCTGCAGGTCACAGTGTGATGATGCTTG 197

RESULT 21
ABX73660
ID ABX73660 standard; DNA; 836 BP.
XX AC ABX73660;
XX DT 18-MAR-2003 (first entry)
XX DE Human novel polynucleotide #488.
XX KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;
XX muscular disorder; respiratory disease; reproductive disorder;
XX gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
XX hyperproliferative disorder; inflammatory disease; allergic reaction;
XX blood related disorder; cancer; immunosuppressive; antiinflammatory;
XX cardiovascular; nephrotoxic; cytostatic; antiallergic; thrombolytic;
XX haemostatic; antiarteriosclerotic.
XX OS Homo sapiens.
XX PN US2002132753-A1.
XX PD 19-SEP-2002.
XX PF 17-JUN-2001; 2001US-0764864.
XX 31-JAN-2000; 2000US-179065P.
XX 04-FEB-2000; 2000US-180628P.
XX 28-JUN-2000; 2000US-214886P.
XX 07-JUL-2000; 2000US-216647P.
XX 07-JUL-2000; 2000US-216880P.
XX 11-JUL-2000; 2000US-217487P.
XX 11-JUL-2000; 2000US-217496P.
XX 14-JUL-2000; 2000US-218290P.
XX 26-JUL-2000; 2000US-220963P.
XX 26-JUL-2000; 2000US-220964P.
XX 14-AUG-2000; 2000US-224518P.
XX 14-AUG-2000; 2000US-224519P.
XX 14-AUG-2000; 2000US-225267P.
XX 14-AUG-2000; 2000US-225268P.
XX 14-AUG-2000; 2000US-225270P.
XX 14-AUG-2000; 2000US-225447P.
XX 14-AUG-2000; 2000US-225757P.
XX 14-AUG-2000; 2000US-225758P.
XX 22-AUG-2000; 2000US-226868P.
XX 30-AUG-2000; 2000US-228924P.
XX 01-SEP-2000; 2000US-229287P.
XX 01-SEP-2000; 2000US-229343P.
XX 01-SEP-2000; 2000US-229344P.
XX 01-SEP-2000; 2000US-229345P.
XX 05-SEP-2000; 2000US-229509P.
XX 05-SEP-2000; 2000US-229513P.
XX 08-SEP-2000; 2000US-231413P.
XX 21-SEP-2000; 2000US-234223P.
XX 21-SEP-2000; 2000US-234274P.
XX 25-SEP-2000; 2000US-234997P.
XX 27-SEP-2000; 2000US-235834P.
XX 29-SEP-2000; 2000US-236327P.
XX 29-SEP-2000; 2000US-236367P.
XX 29-SEP-2000; 2000US-236368P.
XX 29-SEP-2000; 2000US-236369P.

```

```

PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.

XX PA (ROSE/) ROSEN C A.
XX PA (ROBE/) RUBEN S M.
XX PA (BARA/) BARASH S C.
XX FI Rosen CA, Ruben SM, Barash SC;
XX WI; 2003-147444/14.
XX DR P-PSDB; ABUS5400.
XX PT New polypeptides and nucleic acids, useful in gene therapy for
XX treating, inhibiting or preventing e.g. neural, immune system,
XX muscular, respiratory, reproductive, gastrointestinal, pulmonary,
XX cardiovascular or renal disorders -
XX Claim 1; SEQ ID NO 498; 402pp; English.
XX PS

```

The invention relates to human novel polypeptides and their associated polynucleotides. The polypeptides and polynucleotides are useful in gene therapy for treating, inhibiting or preventing neural disorders, immune system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis), muscular disorders, respiratory diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders, gastrointestinal disorders, pulmonary disorders, cardiovascular disorders (e.g. congenital heart defects, Epstein's anomaly and hypoplastic left heart disease), hyperproliferative disorders (e.g. Hodgkin's disease and leukaemia), inflammatory diseases (e.g. septic shock, bursitis and appendicitis), allergic reactions and conditions (e.g. asthma), blood related disorders (e.g. thrombosis, atherosclerosis and myocardial infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent human novel polynucleotides of the invention.

SEQ Sequence 836 BP; 224 A; 179 C; 221 G; 211 T; 1 other;

Query Match 25.2%; Score 32; DB 25; Length 836;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 14 GCCATCTGCAGGTCACAGTGTGATGATGCTTG 45
Db 219 GCCATCTGCAGGTCACAGTGTGATGATGCTTG 250

```

```

RESULT 22
AAC77504
ID AAC77504 standard; cDNA; 962 BP.
XX AC AAC77504;
XX DT 08-FEB-2001 (first entry)
XX DE Human ORFX ORF3059 polynucleotide sequence SEQ ID NO.6117.
XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX vulnerrary; antipariariar; antiparkinsonian; neuroprotective;
XX anticonvulsant; osteopathic; antitarrtritic; immunosuppressant; cardiant;
KW

```

KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antineoplastic;
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antineoplastic disease; coagulation;
 KW thrombosis; contraceptive; ss.
 OS Homo sapiens.
 PN WO200058473-A2.
 XX
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkens RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR P-PSDB; AAB43395.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 PT
 Claim 5; Page 5300; 5507JP; English.
 XX
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
 CC osteoparalytic; anticonvulsant; antiallergic; immunosuppressive;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antineoplastic; antibacterial; antiviral; antifungal; antineoplastic;
 CC antihypertensive; antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy.
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antineoplastic disease, to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX
 SQ Sequence 962 BP; 265 A; 204 C; 243 G; 250 T; 0 other;

Query Match 25.2%; Score 32; DB 21; Length 962;
 Best Local Similarity 100.0%; Pred. No. 4.5e-07;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGTCAGGTCAGTGCCTG 45
 |||||
 DB 202 GCCATCTGCAGGTCAGGTCAGTGCCTG 233

RESULT 23
AAX87320

ID AAX87320 standard; cDNA; 754 BP.
 XX
 AC AAX87320;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE Human sensitive to apoptosis (SAG) gene mutant MM4.
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnary; therapy; mutant; ds.
 KW
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 XX CDS 1..342
 FT /*tag= a
 FT mutation replace(181,T)
 FT /*tag= b
 FT /*note= "C61S mutation"
 XX
 PN WO9932514-A2.
 PD
 PD 01-JUL-1999.
 XX
 PF 15-DEC-1998; 98WO-US26705.
 XX
 PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 XX
 DR WPI; 1999-430152/36.
 DR P-PSDB; AAY06498.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PT
 Claim 15; Page 64-65; 84pp; English.
 XX
 XX This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM4, which codes for a SAG protein (see AAY06498)
 CC in which the Cys residue at position 61 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 50 from TGC to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is at zinc-finger finger 1 of SAG. Single and
 CC double SAG mutants (see AAX87317-11) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM4
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 XX
 SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 23.6%; Score 30; DB 20; Length 754;
 Best Local Similarity 100.0%; Pred. No. 5.2e-06;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGTCAGGTCAGTGCCTG 43
 |||||
 DB 151 GCCATCTGCAGGTCAGGTCAGTGCCTG 180

RESULT 24

```

AAx87329
ID AAX87329 standard; cDNA; 754 BP.
XX
AC AAX87329;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene mutant M13.
XX
KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnerability; therapy; mutant; ds.
XX
XX Homo sapiens.
XX OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..342
FT FT /tag= a
FT FT replace(181,T)
FT FT /tag= b
FT FT /note= "C61S mutation"
FT FT replace(190,T)
FT FT /tag= c
FT FT /note= "C64S mutation"
XX
XX MO9932514-A2.
XX
XX 01-JUL-1999.
XX
XX 15-DEC-1998; 98WO-US26705.
XX
XX 11-SEP-1998; 98US-0099840.
XX 19-DEC-1997; 97US-0068179.
XX
XX (WARN) WARNER LAMBERT CO.
XX
XX Sun Y;
XX WPI: 1999-430152/36.
XX P-PSDB; AAY06507.
XX
XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
XX promoting cell growth and protecting cells against apoptosis
XX
XX Claim 15; Page 79-80; 84pp; English.
XX
XX This is the nucleotide sequence of human sensitive to apoptosis
XX (SAG) mutant gene M13, which codes for a SAG protein (see AAY06507)
XX in which the Cys residues at positions 61 and 64 of the native
XX protein (see AAY06492) are replaced by Ser residues owing to
XX mutations of codons 61 and 64 through site-directed mutagenesis of
XX SAG cDNA. These residues are in zinc ring finger 1 of SAG. Single
XX double SAG mutants (see AAX87317-31) were made in order to determine
XX the role of each cysteine residue of SAG in haem binding and SAG
XX oligomerization. M13 showed greatly reduced haem binding (and a
XX decreased brown colour) but oligomerization was unaffected. SAG is
XX a novel zinc finger protein that promotes cell growth, protects cells
XX from apoptosis, scavenges oxygen radicals and can be used for the
XX reversion of a tumour phenotype. SAG genes, and mutant SAG genes,
XX can be used to protect cells from apoptosis induced by redox
XX reagents. They can also be used for the recombinant production of
XX SAG proteins, which are molecular targets in the development of
XX drugs against neurodegenerative disorders, cancers and muscle
XX dystrophy, and promoting wound healing.
XX
XX Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;
XX
XX Query Match 23.6%; Score 30; DB 20; Length 754;
XX Best Local Similarity 100.0%; Pred. No. 5.2e-06;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX 14 GCCATCTGCAGGCTCCAGTGATGATGCC 43

```

```

DB
151 GCCATCTGCAGGCTCCAGTGATGATGCC 180
|||||
RESULT 25
AAX87315
ID AAX87315 standard; cDNA; 747 BP.
XX
AC AAX87315;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene mutant 1.
XX
KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX neurodegenerative disease; muscular dystrophy; wound healing;
XX vulnerability; therapy; mutant; colon carcinoma; ds.
XX
XX Homo sapiens.
XX OS
XX
FH Key Location/Qualifiers
FT CDS 1..273
FT FT /tag= a
FT FT
XX
XX MO9932514-A2.
XX
XX 01-JUL-1999.
XX
XX 15-DEC-1998; 98WO-US26705.
XX
XX 11-SEP-1998; 98US-0099840.
XX 19-DEC-1997; 97US-0068179.
XX
XX (WARN) WARNER LAMBERT CO.
XX
XX Sun Y;
XX WPI: 1999-430152/36.
XX P-PSDB; AAY06493.
XX
XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
XX promoting cell growth and protecting cells against apoptosis
XX
XX Claim 15; Page 54-55; 84pp; English.
XX
XX This is the nucleotide sequence of deletion mutant 1 of the novel
XX human sensitive to apoptosis gene SAG (see also AAX87314). The cDNA
XX was obtained by PCR amplification of RNA isolated from DLD-1 colon
XX carcinoma cells (ATCC CCL 221). It contains a 7 bp deletion
XX compared with native SAG, starting at nucleotide 170 of the coding
XX region that codes for a potential haem binding site. The frame
XX shift deletion abolishes the downstream zinc finger in the
XX resulting protein (see AAY06493). The mutation was detected by PCR
XX in SAG RNA, but not in genomic DNA. It suggests a possible role
XX for SAG in human carcinogenesis. SAG genes, and mutant SAG
XX genes, can be used to protect cells from apoptosis induced by redox
XX reagents. Antisense SAG genes can also be used to inhibit the growth of
XX tumour cells. The SAG genes can also be used for the recombinant
XX production of the SAG proteins. The SAG proteins can be used to
XX scavenge oxygen radicals in organisms and to promote wound healing.
XX Additionally, the SAG genes or their complements can be used to
XX promote or inhibit the growth of plant cells (all claimed). The SAG
XX protein is also an ideal molecular target in the development of
XX drugs against neurodegenerative disorders, cancers and muscle
XX dystrophy.
XX
XX Sequence 747 BP; 204 A; 155 C; 197 G; 191 T; 0 other;
XX
XX Query Match 22.8%; Score 29; DB 20; Length 747;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-05;
XX Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX 71 CAAGAGACTGTGTGTGTGTGCTGGGAGAG 99

```

Db 201 CAAGAGACTGTGTGTGCTGGGAGA 229

RESULT 26

AAx87318
ID AAX87318 standard; cDNA; 754 BP.

AC AAX87318;

DT 27-SEP-1999 (first entry)

DE Human sensitive to apoptosis (SAG) gene mutant MM2.

KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
KW neurodegenerative disease; muscular dystrophy; wound healing;
KW vulnerability; therapy; mutant; ds.

OS Homo sapiens.
OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..342

FT mutation /*tag= a
replace(157,T)
/*tag= b
/note= "C53S mutation"

PN WO9932514-A2.

PD 01-JUL-1999.

PF 15-DEC-1998; 98WO-US26705.

PR 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.

XX (WARN) WARNER LAMBERT CO.

XX Sun Y;

XX WPI; 1999-430152/36.
XX P-PSDB; AAY06496.

PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
promoting cell growth and protecting cells against apoptosis

XX Claim 15; Page 60-61; 84pp; English.

XX This is the nucleotide sequence of human sensitive to apoptosis
(SAG) mutant gene MM2, which codes for a SAG protein (see AAY06496)
in which the Cys residue at position 53 of the native protein (see
CC AAY06496) is replaced by a Ser residue owing to a mutation of codon
50 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is at a haem binding site of SAG. Single and
CC double SAG mutants (see AAX87317-31) were made in order to determine
the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM2
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversion of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.

XX Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

XX Query Match 22.8%; Score 29; DB 20; Length 754;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-05;
XX Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 71 CAAGAGACTGTGTGTGCTGGGAGA 99
Db 208 CAAGAGACTGTGTGTGCTGGGAGA 236

RESULT 27

AAx87319
ID AAX87319 standard; cDNA; 754 BP.

AC AAX87319;

DT 27-SEP-1999 (first entry)

DE Human sensitive to apoptosis (SAG) gene mutant MM3.

KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
KW neurodegenerative disease; muscular dystrophy; wound healing;
KW vulnerability; therapy; mutant; ds.

OS Homo sapiens.
OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..342

FT mutation /*tag= a
replace(148,T)
/*tag= b
/note= "C50S mutation"
replace(157,T)
/*tag= c
/note= "C53S mutation"

PN WO9932514-A2.

PD 01-JUL-1999.

PF 15-DEC-1998; 98WO-US26705.

PR 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.

XX (WARN) WARNER LAMBERT CO.

XX Sun Y;

XX WPI; 1999-430152/36.
XX P-PSDB; AAY06497.

PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
promoting cell growth and protecting cells against apoptosis

XX Claim 15; Page 62-63; 84pp; English.

XX This is the nucleotide sequence of human sensitive to apoptosis
(SAG) mutant gene MM3, which codes for a SAG protein (see AAY06497)
in which the Cys residues at positions 50 and 53 of the native
CC protein (see AAY06497) are replaced by Ser residues owing to
CC site-directed mutagenesis of SAG cDNA. These residues are at a haem
CC binding site of SAG. Single and double SAG mutants (see AAX87317-31)
CC were made in order to determine the role of each cysteine residue
of SAG in haem binding and SAG oligomerization. MM3 showed greatly
CC reduced haem binding but oligomerization was unaffected. SAG is a
CC novel zinc finger protein that promotes cell growth, protects cells
CC from apoptosis, scavenges oxygen radicals and can be used for the
CC reversion of a tumour phenotype. SAG genes, and mutant SAG genes,
CC can be used to protect cells from apoptosis induced by redox
CC reagents. They can also be used for the recombinant production of
CC SAG proteins, which are molecular targets in the development of
CC drugs against neurodegenerative disorders, cancers and muscle
CC dystrophy, and promoting wound healing.

XX Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;

Query Match 22.8%; Score 29; DB 20; Length 754;
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAAGAGACTGTGTGTGTGTGGGAGA 99
 |||||
 DB 208 CAAGAGACTGTGTGTGTGGGAGA 236

RESULT 28

AAH25847
 ID AAH25847 standard; DNA; 1152 BP.

XX AAH25847;

DT 20-AUG-2001 (first entry)

XX Human apoptosis associated protein 12 coding sequence.

XX Human; apoptosis associated protein 12; SAG protein 12; cancer;

KM haemopathy; HIV infection; immunological disease; inflammation; ds.

XX Homo sapiens.

OS WO200132863-A1.

XX 10-MAY-2001.

XX 30-OCT-2000; 2000MO-CN00406.

XX 29-OCT-1999; 99CN-0119924.

XX (SHAN-) SHANGHAI BIO DOOR GENE TECHNOLOGY LTD.

XX Mao Y, Xie Y;

XX WPI; 2001-335831/35.

DR P-PSDB; AAB98975.

XX Human apoptosis associated protein 12 and encoded polynucleotide;

PT applicable in diagnosis and treatment of malignant tumour, hemopathy,

PT HIV infection; immunological diseases and various inflammation

XX Claim 6; Page 22-23; 27pp; Chinese.

XX The present invention provides the protein and coding sequences of human

CC apoptosis associated (SAG) protein 12. These sequences can be used in

CC the diagnosis and treatment of malignant tumours, haemopathy, HIV

CC infection, immunological diseases and various types of inflammation. The

CC present sequence is the SAG protein 12 coding sequence.

XX Sequence 1152 BP; 292 A; 265 C; 309 G; 286 T; 0 other;

Query Match 22.8%; Score 29; DB 22; Length 1152;

Best Local Similarity 100.0%; Pred. No. 1.7e-05;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAAGAGACTGTGTGTGTGTGGGAGA 99
 |||||
 DB 596 CAAGAGACTGTGTGTGTGGGAGA 624

RESULT 29

ABN40537
 ID ABN40537 standard; DNA; 60 BP.

XX ABN40537;

DT 15-JUN-2002 (first entry)

XX Human spliced transcript detection oligonucleotide SEQ ID NO:13285.

KM Human; mouse; rat; splice transcript; detection; RNA transcript;

XX splice variant; transcriptome; oligonucleotide library; ss.

OS Homo sapiens.

XX WO200210449-A2.

XX 20-JUL-2001; 2001WO-IB01903.

XX 28-JUL-2000; 2000US-221607P.

XX 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257363/30.

XX New oligonucleotide libraries comprising oligonucleotides which

PT selectively hybridize to mRNAs transcribed from a transcription unit of

PT a genome, useful for detecting tissue-, pathology-, and

PT developmental-specific genes

PS Example 1; SEQ ID 13285; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting

CC messenger RNAs that populate a (sub-)transcriptome, where the

CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple

CC transcription units that populate a genome. The library comprises

CC several oligonucleotides, each capable of hybridizing selectively to a

CC set of messenger RNAs transcribed from a given transcription unit of

CC the genome, which encodes one or more messenger RNA splice variants.

CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or

CC quantitatively characterizing the corresponding transcriptome, and in

CC detecting RNA transcripts and splice variants of human or animal

CC transcriptomes. The libraries may also be used as specialised mini

CC libraries to detect transcripts of a sub-transcriptome under a

CC particular biological or pathological state, and so allowing the

CC detection of tissue- and pathology-specific genes such as those genes

CC only expressed in specific tissue under a specific pathological

CC condition; to detect developmental specific genes; and to detect RNA

CC transcripts and splice variants of a transcriptome of a patient suffering

CC from a particular disorder. ABN27253 to ABN59589 represent

CC oligonucleotide sequences from rats, humans and mice, which are used in

CC the exemplification of the present invention.

CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 60 BP; 10 A; 13 C; 25 G; 12 T; 0 other;

Query Match 19.7%; Score 25; DB 24; Length 60;

Best Local Similarity 100.0%; Pred. No. 0.0023;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGAGTCCAGGTGATG 38
 |||||
 DB 36 GCCATCTGCAGAGTCCAGGTGATG 60

RESULT 30

AB211414
 ID AB211414 standard; cDNA; 439 BP.

XX AB211414;

DT 20-JAN-2003 (first entry)

XX Human polynucleotide SEQ ID NO 296.

KW	Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW	cell-proliferative disorder; neurodegenerative disease; bacterial;
KW	Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW	multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW	atrial fibrillation; cytoskeletal; immunomodulator; neurotrophic; neuroprotective;
KW	antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW	haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW	antiarthritic; gene; ss.
XX	
OS	Homo sapiens.
PN	WO200270539-A2.
XX	
PD	12-SEP-2002.
XX	
PF	05-MAR-2002; 2002WO-US05095.
XX	
PR	05-MAR-2001; 2001US-0799451.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QH, Ren F;
PI	Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI	Wehrman T, Wang J, Wang D, Drmanac RT;
XX	
DR	WPI: 2002-759812/82.
XX	P-PSDB; ABP69197.
PT	New polynucleotides comprising sequences assembled from expressed
PT	sequence tags (ESTs), useful for treating cell-proliferative,
PT	neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
PT	platelet or coagulation disorders -
XX	
PS	Claim 1; SEQ ID NO 296; 1012bp + Sequence listing; English.
XX	
CC	The invention relates to an isolated polynucleotide (1) comprising a
CC	nucleotide sequence selected from any of 948 sequences
CC	(AB21119-AB212066) or their mature protein coding portion, active domain
CC	coding protein or complementary sequences. The polynucleotides are useful
CC	for identifying expressed genes or for physical mapping of human genome.
CC	The encoded polypeptides (ABP6902-ABP69849) are useful as molecular
CC	weight markers, as a food supplement, for generating antibodies, in
CC	medical imaging, screening and diagnostic assays and for treating
CC	cell-proliferative disorders (cancer), neurodegenerative diseases
CC	(Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
CC	sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
CC	disorders, platelet or coagulation disorders, wound, burns, incision,
CC	ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
CC	parasitic), arthritis, etc.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SO	Sequence 439 BP; 69 A; 142 C; 162 G; 66 T; 0 other;
	Query Match 19.7%; Score 25; DB 24; Length 439;
	Best Local Similarity 100.0%; Pred.No.0.0023;
	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	14 GCCATCTGCAGGGTCCAGGTGATGG 38
DB	231 GCCATCTGCAGGGTCCAGGTGATGG 255
RESULT 31	
AACT7493	AACT7493 standard; cDNA; 441 BP.
XX	
AC	AACT7493;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Human OREF3048 polynucleotide sequence SEQ ID NO:6095.

XX Human, open reading frame, ORFX, detection, cytosolic, hepatotropic;
 KM vulnerability; antiparasitic; antiparkinsonian; nootropic; neuroprotective;
 KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KM antiviral; antibacterial; antifungal; antineumatic; antithyroid;
 KM antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KM thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 XX
 PR 02-APR-1999; 99US-0127636.
 XX
 PR 05-APR-1999; 99US-0127728.
 XX
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CUBA-) CUBAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 XX
 DR P-PSDB: AAB43284.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 5; Page 5279; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosolic; hepatotropic; vulnerability;
 CC antiparasitic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 SO Sequence 441 BP; 98 A; 113 C; 139 G; 90 T; 1 other;
 QY
 Query Match 19.7%; Score 25; DB 21; Length 441;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 14 GCCATCTGCAGGGCTCCAGGTCATGG 38
 |||||
 DB 184 GCCATCTGCAGGGCTCCAGGTCATGG 208


```

RESULT 32
AA87316
ID   AA87316 standard; cDNA; 706 BP.
XX
XX
AC   AA87316;
XX
DT   27-SEP-1999 (first entry)
XX
XX
DE   Human sensitive to apoptosis (SAG) gene mutant 2.
XX
XX
KW   SAG gene; sensitive to apoptosis; human; cancer; tumour;
XX   neurodegenerative disease; muscular dystrophy; wound healing;
XX   vulnery; therapy; mutant; colon carcinoma; ds.
XX
OS   Homo sapiens.
XX
FH   Key Location/Qualifiers
FT   CDS 1..294
FT   /tag= a
XX
XX
PN   MO9932514-A2.
XX
XX
PD   01-JUL-1999.
XX
XX
PF   15-DEC-1998; 98WO-US26705.
XX
XX
PR   11-SEP-1998; 98US-009840.
XX   19-DEC-1997; 97US-0068179.
XX
XX
PA   (WARN ) WARNER LAMBERT CO.
XX
XX
PI   Sun Y;
XX
DR   MPI; 1999-430152/36.
XX
DR   P-PSDB; AA06494.
XX
XX
PT   SAG: Sensitive to Apoptosis Gene and related proteins, useful for
XX   promoting cell growth and protecting cells against apoptosis
XX
XX
PS   Claim 15; Page 55-56; 84pp; English.
XX
XX
CC   This is the nucleotide sequence of deletion mutant 2 of the novel
XX   human sensitive to apoptosis gene SAG (see also AA87314). The cDNA
XX   was obtained by PCR amplification of RNA isolated from DLD-1 colon
XX   carcinoma cells (ATCC CCL 221). It contains a 48 bp deletion
XX   compared with native SAG, starting at nucleotide 177 of the coding
XX   region that codes for a potential haem binding site. The in-frame
XX   deletion eliminates 16 amino acids in the encoded protein (see
XX   AA06494) but retains the zinc finger motif. The mutation was
XX   detected by PCR in SAG RNA, but not in genomic DNA. It suggests a
XX   possible role for SAG in human carcinogenesis. SAG genes, and
XX   mutant SAG genes, can be used to protect cells from apoptosis
XX   induced by redox reagents. Antisense SAG genes can be used to
XX   inhibit the growth of tumour cells. The SAG genes can also be used
XX   for the recombinant production of the SAG proteins. The SAG
XX   proteins can be used to scavenge oxygen radicals in organisms and
XX   to promote wound healing. Additionally, the SAG genes or their
XX   complements can be used to promote or inhibit the growth of plant
XX   cells (all claimed). The SAG protein is also an ideal molecular
XX   target in the development of drugs against neurodegenerative
XX   disorders, cancers and muscle dystrophy.
XX
XX
SQ   Sequence 706 BP; 189 A; 147 C; 189 G; 181 T; 0 other;
XX
Query Match 19.7%; Score 25; DB 20; Length 706;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 GCCATCTGCAGGGTCCAGGTGATGG 38
DB 151 GCCATCTGCAGGGTCCAGGTGATGG 175

```

```

RESULT 33
AAL16209
ID   AAL16209 standard; cDNA; 264 BP.
XX
XX
AC   AAL16209;
XX
DT   07-DEC-2001 (first entry)
XX
XX
DE   Human breast cancer expressed polynucleotide 8666.
XX
XX
KW   Human; breast cancer; cell marker; cytostatic; ss.
XX
XX
OS   Homo sapiens.
XX
XX
PN   WO200151628-A2.
XX
XX
PD   19-JUL-2001.
XX
XX
PF   10-JAN-2001; 2001WO-US00798.
XX
XX
PR   14-JAN-2000; 2000US-0176077.
XX   14-MAR-2000; 2000US-0189167.
XX   24-MAR-2000; 2000US-0192099.
XX   29-MAR-2000; 2000US-0193480.
XX   15-MAY-2000; 2000US-0205230.
XX   09-JUN-2000; 2000US-0211315.
XX   25-JUL-2000; 2000US-0220534.
XX
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
PI   Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX
DR   MPI; 2001-451856/48.
XX
XX
PT   New peptide useful as a marker for the diagnosis of breast cancer
XX
XX
PS   Claim 1; Page 1568; 3695pp; English.
XX
XX
CC   The invention relates to human breast cancer expressed polynucleotides
XX   (AAL07544-AAL26789) and methods of assessing whether a patient is
XX   afflicted with breast cancer by examining the correlation between the
XX   expression of certain markers and the cancerous state of breast cells.
XX   The polynucleotides and encoded polypeptides are potential markers for
XX   detecting, diagnosing, monitoring, characterizing treating and
XX   potentially preventing breast cancer. The polynucleotides and encoded
XX   polypeptides are also useful for isolating compounds with cytostatic
XX   activity.
XX
XX
SQ   Sequence 264 BP; 64 A; 59 C; 74 G; 62 T; 5 other;
XX
Query Match 16.5%; Score 21; DB 22; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 107 CATTCTTCACAACTGCTGC 127
DB 57 CATTCTTCACAACTGCTGC 77
XX
XX
RESULT 34
AAL25052
ID   AAL25052 standard; cDNA; 596 BP.
XX
XX
AC   AAL25052;
XX
DT   07-DEC-2001 (first entry)
XX
XX
DE   Human breast cancer expressed polynucleotide 17509.
XX
XX
KW   Human; breast cancer; cell marker; cytostatic; ss.
XX
XX

```

OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer -
XX
PS Claim 1; Page 3235; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AA07544-AA026789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 596 BP; 166 A; 124 C; 144 G; 162 T; 0 other;
XX
SQ
XX
Query Match 16.5%; Score 21; DB 22; Length 596;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 107 CATTCCTCCACAACGTCTGC 127
DB 14 CATTCCTCCACAACGTCTGC 34
XX
RESULT 35
AAL22239
ID AAL22239 standard; cDNA; 630 BP.
XX
AC AAL22239;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 14696.
XX
KM Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR

PR 25-JUL-2000; 2000US-0220534.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer -
XX
PS Claim 1; Page 2649-2650; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AA07544-AA026789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 630 BP; 213 A; 125 C; 121 G; 171 T; 0 other;
XX
SQ
XX
Query Match 16.5%; Score 21; DB 22; Length 630;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 107 CATTCCTCCACAACGTCTGC 127
DB 46 CATTCCTCCACAACGTCTGC 66
XX
RESULT 36
ABQ90662
ID ABQ90662 standard; DNA; 225 BP.
XX
AC ABQ90662;
XX
DT 01-OCT-2002 (first entry)
XX
DE M. capsulatus gene #647 for DNA array.
XX
KM Micro array; gene; ds; differential expression; gene expression.
XX
OS Methylococcus capsulatus.
XX
PN WO200255655-A2.
XX
PD 18-JUL-2002.
XX
PF 14-JAN-2002; 2002WO-NO00019.
XX
PR 12-JAN-2001; 2001NO-0000235.
PR 12-JAN-2001; 2001NO-0000239.
XX
PA (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
XX
PI (TIGR-) TIGR.
XX
PI Birkeleand NK, Eichhammer I, Jonassen I, Jensen HB, Lien T;
PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;
PI Salzberg SL;
XX
DR WPI; 2002-557818/59.
XX
PT Novel DNA array useful for determining differential expression of
PT Methylococcus capsulatus genes, comprises polynucleotides or
PT oligonucleotides representative for a selective number of Methylococcus
PT capsulatus genes -
XX
PS Claim 14; Page 318; 678pp; English.
XX
CC The invention relates to a novel DNA array giving a representation of a

CC number of *Methylococcus capsulatus* genes. The method of the invention is
CC useful for determination of the differential expression of the genes of
CC *M. capsulatus*, and for studying gene expression on a genomic scale and in
CC gene expression assays of *M. capsulatus* genes. The sequences shown in
CC ABQ9016-ABQ9185 represent *M. capsulatus* genes for use in arrays of the
CC invention.
XX
SQ Sequence 225 BP; 56 A; 42 C; 72 G; 55 T; 0 other;
Query Match 15.0%; Score 19; DB 24; Length 225;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 TACCTGTGCATCTGCAGG 25
Db 31 TACCTGTGCATCTGCAGG 49
RESULT 37
AAL36982
ID AAL36982 standard; DNA; 498 BP.
XX
AC AAL36982;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3347.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0250150.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-02559678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451937/48.
XX
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
XX Example 2; SEQ ID NO 3347; 781bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 498 BP; 115 A; 115 C; 109 G; 159 T; 0 other;
XX
XX Query Match 15.0%; Score 19; DB 22; Length 498;
XX Best Local Similarity 100.0%; Pred. No. 3.4;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 AACACGACAGGACTCT 82
Db 403 AACACGACAGGACTCT 421
RESULT 38
ID ABX59970 standard; cDNA; 498 BP.
ABX59970
AC ABX59970;
XX
XX 26-FEB-2003 (first entry)
XX
XX cDNA encoding novel human musculoskeletal system antigen #2314.
XX
XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
XX post-operative tissue repair; limb regeneration; neuronal growth;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX AIDS-related complex; chondrocyte growth; bone regeneration;
XX periodontal regeneration; tissue transport; bone graft; skin aging;
XX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
XX cell growth; organ transplant; cell differentiation; body height;
XX weight; hair colour; eye colour; skin; percentage of adipose tissue;
XX pigmentation; cosmetic surgery; metabolism; biorhythm; circadian rhythm;
XX depression; tendency for violence; pain; reproductive capability;
XX hormone level; endocrine level; appetite; libido; memory; stress;
XX storage capability; fat content; lipid content; protein content;
XX carbohydrate content; vitamin content; cofactor content;
XX nutritional component.
OS Homo sapiens.
XX
XX US2002147140-A1.
XX
XX 10-OCT-2002.
XX
XX 17-JAN-2001; 2001US-0764877.
XX
XX 31-JAN-2000; 2000US-179065P.
XX 04-FEB-2000; 2000US-180628P.
XX 28-JUN-2000; 2000US-214886P.
XX 07-JUL-2000; 2000US-216647P.
XX 07-JUL-2000; 2000US-216880P.
XX 11-JUL-2000; 2000US-217487P.
XX 14-JUL-2000; 2000US-218290P.
XX 26-JUL-2000; 2000US-220963P.
XX 26-JUL-2000; 2000US-220964P.
XX 14-AUG-2000; 2000US-224518P.
XX 14-AUG-2000; 2000US-224519P.
XX 14-AUG-2000; 2000US-225267P.
XX 14-AUG-2000; 2000US-225268P.
XX 14-AUG-2000; 2000US-225270P.
XX 14-AUG-2000; 2000US-225470P.
XX 14-AUG-2000; 2000US-225757P.
XX 14-AUG-2000; 2000US-225758P.
XX 22-AUG-2000; 2000US-226868P.
XX 30-AUG-2000; 2000US-228924P.
XX 01-SEP-2000; 2000US-229287P.
XX 01-SEP-2000; 2000US-229287P.
XX 01-SEP-2000; 2000US-229343P.
XX 01-SEP-2000; 2000US-229344P.
XX 01-SEP-2000; 2000US-229345P.
XX 05-SEP-2000; 2000US-229509P.
XX 05-SEP-2000; 2000US-229513P.
XX 08-SEP-2000; 2000US-231413P.
XX 21-SEP-2000; 2000US-234223P.
XX 21-SEP-2000; 2000US-234274P.
XX 25-SEP-2000; 2000US-234997P.
XX 27-SEP-2000; 2000US-235834P.
XX 29-SEP-2000; 2000US-236327P.
XX 29-SEP-2000; 2000US-236367P.

PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239355P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.

PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.

PI Rosen CA, Ruben SM, Barash SC;

DR WPI; 2003-128199/12.

PT Isolated nucleic acid molecules encoding musculoskeletal system
 associated polypeptides, useful for detecting disorders, e.g., cancer -
 PS Disclosure; SEQ ID NO 3347; 321pp; English.

CC The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid; stimulates re-vascularisation of ischemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC and other cardiovascular conditions; treats wounds due to injuries;
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryos; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells; besides, haematopoietic lineage; modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
 CC mammal's metal state or physical state by influencing biorhythms;
 CC cardiac rhythms; depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory, or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC or other nutritional components. This sequence encodes a novel human
 CC musculoskeletal system antigen.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.

XX Sequence 498 BP; 115 A; 115 C; 109 G; 159 T; 0 other;

XX Query Match 15.0%; Score 19; DB 25; Length 498;
 XX Best Local Similarity 100.0%; Pred. No. 3.4;
 XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 AACACAGAGAGACTGT 82
 DB 403 AACACAGAGAGACTGT 421

RESULT 39
 ID AA87337
 ID AA87337 standard; DNA; 18 BP.

XX AA87337;

DT 27-SEP-1999 (first entry)

DE Human sensitive to apoptosis gene (SAG) primer hSAG.M1.

XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnerability; therapy; PCR; primer; ss.

XX Synthetic.
 OS Homo sapiens.

XX WO9932514-A2.

XX 01-JUL-1999.

XX 15-DEC-1998; 98WO-US26705.

XX 11-SEP-1998; 98US-0099840.

XX 19-DEC-1997; 97US-0068179.

XX (WARN) WARNER LAMBERT CO.

XX Sun Y;

XX WPI; 1999-430152/36.

PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PS Example 13; Page 26; 84pp; English.

CC Primer hSAG.M1 corresponds to nucleotides 151-171 of the human
 CC sensitive to apoptosis gene (SAG) cDNA clone provided in AA87314.
 CC It was used with primer SAGT.02-1 (see AA87338) in the RT-PCR
 CC amplification of RNA isolated from 20 human tumour lines and
 CC transformed lines originating from lung, brain, kidney, prostate,
 CC testis, nasopharynx, bone, cervix and foreskin. 2 SAG deletion
 CC mutants (see AA87315-16) were detected in cancer cell lines
 CC originating from colon and testis, suggesting a possible role for
 CC SAG in human carcinogenesis. SAG (see also AA87314) is a
 CC redox-sensitive, haem-binding protein that promotes cell growth,
 CC protects cells from apoptosis, and scavenges oxygen radicals. It
 CC can be used to reverse a tumour phenotype.

XX Sequence 18 BP; 3 A; 6 C; 6 G; 3 T; 0 other;

XX Query Match 14.2%; Score 18; DB 20; Length 18;
 XX Best Local Similarity 100.0%; Pred. No. 12;
 XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 GCCATTCGACGGTCCAG 31
 DB 1 GCCATTCGACGGTCCAG 18

RESULT 40

XX AA84630/C

XX ID AA84630 standard; cDNA; 239 BP.

XX AA84630;
 XX 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #20434.
XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI WPI; 2001-639362/73.
XX DR P-PSDB; ABG20443.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 1; SEQ ID NO 20434; 103pp; English.
XX XX
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_poc_sequences.
XX XX
XX SQ Sequence 239 BP; 47 A; 87 C; 58 G; 47 T; 0 other;
XX XX
XX Query Match 13.4%; Score 17; DB 23; Length 239;
XX Best Local Similarity 100.0%; Pred. No. 39;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX QY 15 CCATCTGCAGGGTCCAG 31
XX Db 139 CCATCTGCAGGGTCCAG 123
XX XX
XX RESULT 41
XX ID ABA20540/c
XX XX ABA20540 standard; cDNA; 348 BP.
XX AC ABA20540;
XX XX
XX DT 10-FEB-2003 (first entry)
XX XX
XX DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #2597.

XX XX
XX KM Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
XX KM cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
XX KM arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
XX KM complex carbohydrate; gene replacement therapy; immunosuppressive;
XX KM antiinflammatory; antibacterial; antithyroid; cerebroprotective;
XX KM antiasthmatic; vasotropic.
XX OS Homo sapiens.
XX PN US2002110548-A1.
XX PD 15-AUG-2002.
XX PF 11-JUN-2001; 2001US-0878574.
XX PR 22-NOV-1996; 96US-0753233.
XX PR 03-DEC-1997; 97US-0984246.
XX PR 09-SEP-1998; 98US-0149674.
XX PR 14-JUN-1999; 99US-0333177.
XX PA (GENV) GENETICS INST INC.
XX PI Sullivan F, Kriz R, Kumar R;
XX PI WPI; 2003-066673/06.
XX DR
XX PT New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
XX PT peptide, for manufacturing complex carbohydrates, or as targets for
XX PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant
XX PT rejection -
XX PS Disclosure; SEQ ID NO 2599; 6pp; English.
XX XX
XX CC The invention relates to a composition comprising a human GDP-mannose
XX CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
XX CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
XX CC in a mammalian subject and for treating or ameliorating diseases affected
XX CC by the level of cellular fucosylation or diseases affected by the
XX CC fucosylation of glycoconjugates. These diseases include arthritis,
XX CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or
XX CC infection. The GM4,6D peptide or a polynucleotide encoding it is also
XX CC useful for manufacturing complex carbohydrates and as targets for
XX CC screening small molecule antagonists of the activity of the enzyme. The
XX CC polynucleotide is useful in developing an assay for defects in the
XX CC enzyme, as well as in gene replacement therapy. Sequences
XX CC ABA17942-ABA17944 and ABA17947-ABA33716 represent DNA molecules encoding
XX CC human GM4,6D peptides of the invention.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from USPTO
XX CC at seqdata.uspto.gov/sequence.html.
XX XX
XX SQ Sequence 348 BP; 114 A; 56 C; 75 G; 103 T; 0 other;
XX XX
XX Query Match 13.4%; Score 17; DB 25; Length 348;
XX Best Local Similarity 100.0%; Pred. No. 39;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX QY 101 TGTATCAATTCCTTCCA 117
XX Db 205 TGTATCAATTCCTTCCA 189
XX XX
XX RESULT 42
XX ID ABA46329
XX XX ABA46329 standard; DNA; 355 BP.
XX AC ABA46329;
XX XX
XX DT 01-FEB-2002 (first entry)
XX XX
XX DE Human breast cell single exon nucleic acid probe #5024.

```

KM Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer; ss.
XX
XX Homo sapiens.
OS
XX WO200157271-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JUN-2001; 2001WO-US00662.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX DR WPI; 2001-496933/54.
XX
XX PT New spatially-addressable set of single exon nucleic acid probes,
XX PT useful for measuring gene expression in sample derived from human
XX PT breast, comprises number of single exon nucleic acid probes
XX
XX PS Claim 1; SEQ ID NO 5024; 327bp + sequence listing; English.
XX
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting
XX CC the probes with a collection of detectably labelled nucleic acids
XX CC derived from mRNA of human breast, and then measuring the label
XX CC bound to each probe of the microarray. The probes are useful for
XX CC verifying the expression of regions of genomic DNA predicted to
XX CC encode proteins. They are useful for gene discovery, and for
XX CC determining predisposition and/or prognosing breast disease. Gene
XX CC expression analysis is useful for assessing the toxicity of chemical
XX CC agents on cells. The microarray of this invention presents a far greater
XX CC diversity of probes for measuring gene expression, with far less bias
XX CC than expressed sequence tag microarrays. The method is suitable for
XX CC rapid production of functional information from genomic sequence. The
XX CC present sequence is a single exon nucleic acid probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;
XX
XX Query Match 13.4%; Score 17; DB 22; Length 355;
XX Best Local Similarity 100.0%; Pred. No. 39;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 25 GGTCCAGGTGATGATG 41
DB 281 GGTCCAGGTGATGATG 297

```

```

XX
XX OS Homo sapiens.
XX
XX PN WO200157271-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JUN-2001; 2001WO-US00662.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX DR WPI; 2001-496933/54.
XX
XX PT New spatially-addressable set of single exon nucleic acid probes,
XX PT useful for measuring gene expression in sample derived from human
XX PT breast, comprises number of single exon nucleic acid probes
XX
XX PS Claim 4; SEQ ID NO 10126; 327bp + sequence listing; English.
XX
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting
XX CC the probes with a collection of detectably labelled nucleic acids
XX CC derived from mRNA of human breast, and then measuring the label
XX CC bound to each probe of the microarray. The probes are useful for
XX CC verifying the expression of regions of genomic DNA predicted to
XX CC encode proteins. They are useful for gene discovery, and for
XX CC determining predisposition and/or prognosing breast disease. Gene
XX CC expression analysis is useful for assessing the toxicity of chemical
XX CC agents on cells. The microarray of this invention presents a far greater
XX CC diversity of probes for measuring gene expression, with far less bias
XX CC than expressed sequence tag microarrays. The method is suitable for
XX CC rapid production of functional information from genomic sequence. The
XX CC present sequence is a single exon nucleic acid probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;
XX
XX Query Match 13.4%; Score 17; DB 22; Length 355;
XX Best Local Similarity 100.0%; Pred. No. 39;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 25 GGTCCAGGTGATGATG 41
DB 281 GGTCCAGGTGATGATG 297

```

```

RESULT 43
ABAS1431
ID ABAS1431 standard; DNA; 355 BP.
XX
XX AC ABAS1431;
XX
XX DT 01-FEB-2002 (first entry)
XX
XX DE Human breast cell single exon nucleic acid probe #10126.
XX
XX KW Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer; ss.

```

```

RESULT 44
ABAS6889
ID ABAS6889 standard; DNA; 355 BP.
XX
XX AC ABAS6889;
XX
XX DT 01-FEB-2002 (first entry)
XX
XX DE Human foetal liver single exon nucleic acid probe #5194.
XX
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX OS Homo sapiens.
XX

```

```
PN WO200157277-A2.
XX
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00669.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483447/52.
XX
XX PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX PS Claim 1; SEQ ID NO 5194; 639bp + sequence listing; English.
XX
XX CC The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;

Query Match 13.4%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred.No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GGTCGAGTGATGATG 41
DB 281 GGTCGAGTGATGATG 297

RESULT 45
ABA69454
ID ABA69454 standard; DNA; 355 BP.
XX
XX ABA69454;
XX
XX 01-FEB-2002 (first entry)
XX
XX DE Human foetal liver single exon nucleic acid probe #17759.
XX
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157277-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00669.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
```

```
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483447/52.
XX
XX PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX PS Claim 4; SEQ ID NO 17759; 639bp + sequence listing; English.
XX
XX CC The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;

Query Match 13.4%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred.No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GGTCGAGTGATGATG 41
DB 281 GGTCGAGTGATGATG 297

RESULT 46
ABA26508
ID ABA26508 standard; DNA; 355 BP.
XX
XX ABA26508;
XX
XX 23-JAN-2002 (first entry)
XX
XX DE Probe #4974 for gene expression analysis in human heart cell sample.
XX
XX KW Human; gene expression; heart; microarray; vascular system; probe;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157274-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00666.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488899/53.
XX
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX PS Claim 1; SEQ ID NO 4974; 530bp; English.
```


XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;

Query Match 13.4%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GGTCCAGGTGATGATG 41
DB 281 GGTCCAGGTGATGATG 297

RESULT 47
ABA36383
ID ABA36383 standard; DNA; 355 BP.
XX ABA36383;
XX
XX
XX 23-JAN-2002 (first entry)
XX
DE Probe #14849 for gene expression analysis in human heart cell sample.
XX
XX Human; gene expression; heart; microarray; vascular system; probe;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 4; SEQ ID No 14849; 530pp; English.

The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system

CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;

Query Match 13.4%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GGTCCAGGTGATGATG 41
DB 281 GGTCCAGGTGATGATG 297

RESULT 48
AAK04990
ID AAK04990 standard; DNA; 355 BP.
XX AAK04990;
XX
XX
XX 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 4981.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 4981; 650pp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.

Query Match 13.4%; Score 17; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GGTCCAGGTGATGATG 41

DB 281 GGTCGAGTGATGATG 297

RESULT 49

ID AAK17713 standard; DNA; 355 BP.

XX AAK17713;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 17704.

XX Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX epilepsy; cancer; ss.

OS Homo sapiens.

PN MO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

PT Single exon nucleic acid probes for analyzing gene expression in human

PS Example 4; SEQ ID NO: 17704; 650bp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

CC invention.

XX Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;

SQ Query Match 13.4%; Score 17; DB 22; Length 355;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GGTCGAGTGATGATG 41
 DB 281 GGTCGAGTGATGATG 297

RESULT 50

ID AAK30523 standard; DNA; 355 BP.

XX AAK30523;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 5080.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for

PS analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 5080; 658bp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is one of

CC the probes of the invention.

XX Sequence 355 BP; 122 A; 47 C; 109 G; 77 T; 0 other;

SQ Query Match 13.4%; Score 17; DB 22; Length 355;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GGTCGAGTGATGATG 41
 DB 281 GGTCGAGTGATGATG 297

Search completed: November 7, 2003, 11:58:13

UOD time : 131.53 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 05:52:38 ; Search time 34.4064 Seconds
(without alignments)
1629.223 Million cell updates/sec

Title: US-09-509-779-1_COPY_154_280

Perfect score: 127
Sequence: 1 GTGCATACCTGTCATCT.....ATTCTTCACACATGCTGCTC 127

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued Patents_NA:*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	13.4	363	4	US-09-328-352-1220
2	17	13.4	942	4	US-09-252-991A-11847
3	17	13.4	978	4	US-09-252-991A-11595
4	17	13.4	1605	4	US-09-252-991A-11780
5	17	13.4	3494	3	US-09-139-802-200
6	17	13.4	3494	4	US-09-659-786-200
7	17	13.4	8654	4	US-08-961-527-98
8	16	12.6	475	4	US-09-149-476-239
9	16	12.6	1382	2	US-08-596-387B-123
10	16	12.6	1382	4	US-09-067-615-123
11	16	12.6	1382	4	US-08-596-387B-123
12	16	12.6	1382	4	PCT-US95-09816A-123
13	16	12.6	1385	2	US-08-596-387B-121
14	16	12.6	1385	4	US-09-067-615-121
15	16	12.6	1385	5	PCT-US95-09816A-121
16	16	12.6	1508	2	US-08-596-387B-122
17	16	12.6	1508	3	US-08-960-190A-24
18	16	12.6	1508	4	US-09-067-615-122
19	16	12.6	1508	5	PCT-US95-09816A-122
20	16	12.6	3208	4	US-09-780-016-27
21	16	12.6	4713	4	US-09-194-285-7
22	16	12.6	5769	1	US-08-652-971-1
23	16	12.6	5769	2	US-08-991-258A-1
24	16	12.6	5769	2	US-08-769-399-1
25	16	12.6	5769	3	US-08-991-953A-1
26	15	11.8	196	3	US-08-663-082-1
27	15	11.8	219	4	US-09-328-352-2677

C	28	15	11.8	266	4	US-09-313-294A-3641	Sequence 3641, Ap
	29	15	11.8	266	4	US-09-702-705-1559	Sequence 1559, Ap
	30	15	11.8	266	4	US-09-736-457-1559	Sequence 1559, Ap
C	31	15	11.8	275	4	US-09-313-294A-2588	Sequence 2588, Ap
	32	15	11.8	318	4	US-09-016-434-1031	Sequence 1031, Ap
	33	15	11.8	474	1	US-08-412-614-93	Sequence 93, Appl
	34	15	11.8	474	2	US-08-635-761-93	Sequence 93, Appl
	35	15	11.8	474	3	US-09-312-520-93	Sequence 94, Appl
	36	15	11.8	476	1	US-08-412-614-94	Sequence 94, Appl
	37	15	11.8	476	2	US-08-635-761-94	Sequence 94, Appl
	38	15	11.8	476	3	US-09-312-520-94	Sequence 94, Appl
	39	15	11.8	819	1	US-08-309-182B-2	Sequence 2, Appl
	40	15	11.8	1426	3	US-09-121-425-3	Sequence 3, Appl
C	41	15	11.8	1426	4	US-09-634-493A-3	Sequence 3, Appl
	42	15	11.8	2238	1	US-08-389-668A-1	Sequence 1, Appl
C	43	15	11.8	2238	1	US-08-732-506-1	Sequence 1, Appl
C	44	15	11.8	2238	5	PCT-US95-05768-1	Sequence 1, Appl
C	45	15	11.8	2508	1	US-07-959-943-10	Sequence 10, Appl

ALIGNMENTS

```
RESULT 1
US-09-328-352-1220
; Sequence 1220, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1220
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1220

Query Match          13.4%; Score 17; DB 4; Length 363;
Best local similarity 100.0%; Pred. No. 5.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      80 TGTGTTGCTGCGG 96
      |||||
Db      213 TGTGTTGCTGCGG 229

RESULT 2
US-09-252-991A-11847
; Sequence 11847, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,768
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11847
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11847

Query Match          13.4%; Score 17; DB 4; Length 942;
```

Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 GGTGATGATGCTGCC 47
Db 489 GGTGATGATGCTGCC 505

RESULT 3

US-09-252-991A-11595/c
; Sequence 11595, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11595
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11595

Query Match 13.4%; Score 17; DB 4; Length 978;
Best Local Similarity 100.0%; Pred. No. 5.9;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 GGTGATGATGCTGCC 47
Db 424 GGTGATGATGCTGCC 408

RESULT 4

US-09-252-991A-11780
; Sequence 11780, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11780
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11780

Query Match 13.4%; Score 17; DB 4; Length 1605;
Best Local Similarity 100.0%; Pred. No. 5.9;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 GGTGATGATGCTGCC 47
Db 409 GGTGATGATGCTGCC 425

RESULT 5

US-09-139-802-200/c
; Sequence 200, Application US/09139802

Patent No. 6180084
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; FILE REFERENCE: P-LJ 3203
; CURRENT APPLICATION NUMBER: US/09/139,802
; PRIOR FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 08/926,914
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 08/710,067
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 200
; LENGTH: 3494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(3024)
US-09-139-802-200

Query Match 13.4%; Score 17; DB 3; Length 3494;
Best Local Similarity 100.0%; Pred. No. 6;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 CCATCTGAGGGTCCAG 31
Db 1762 CCATCTGAGGGTCCAG 1746

RESULT 6

US-09-659-786-200/c
; Sequence 200, Application US/09659786
; Patent No. 6491894
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; FILE REFERENCE: P-LJ 3203
; CURRENT APPLICATION NUMBER: US/09/659,786
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: 08/926,914
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 08/710,067
; PRIOR FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 200
; LENGTH: 3494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(3024)
US-09-659-786-200

Query Match 13.4%; Score 17; DB 4; Length 3494;
Best Local Similarity 100.0%; Pred. No. 6;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 CCATCTGAGGGTCCAG 31
Db 1762 CCATCTGAGGGTCCAG 1746

RESULT 7

US-08-961-527-98/c

Sequence 98, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 8654 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-98

Query Match 13.4%; Score 17; DB 4; Length 8654;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GACTGTGTGTGCTG 93
Db 1105 GACTGTGTGTGCTG 1089

RESULT 8
US-09-149-476-239/C
Sequence 239, Application US/09149476
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
EARLIER FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 12.6%; Score 16; DB 4; Length 458;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 18 TCTGAGGTCACGCT 33
Db 238 TCTGAGGTCACGCT 223

RESULT 9
US-09-149-476-76/c
Sequence 76, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597

[illegible]

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 12.6%; Score 16; DB 4; Length 475;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TCTGCGGTCGAGT 33

Db 226 TCTGCGGTCGAGT 211

RESULT 10
US-08-596-387B-123
Sequence 123, Application US/08596387B
Patent No. 5869270
GENERAL INFORMATION:
APPLICANT: Rhode, Peter R.
APPLICANT: Jiao, Jin-An
APPLICANT: Burkhardt, Martin
APPLICANT: Wong, Hing
TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dade International, Inc.
STREET: 1717 Deerfield Road
CITY: Deerfield
STATE: Illinois
COUNTRY: USA
ZIP: 60015
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/596,387B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09816
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/382,454
FILING DATE: 01-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,302
FILING DATE: 29-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pearson, Louise S.

REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STR-4665-CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 267-5300
TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 123:
SEQUENCE CHARACTERISTICS:
LENGTH: 1382 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 6..1382
US-08-596-387B-123

Query Match 12.6%; Score 16; DB 2; Length 1382;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 105 ACCATCTCTCCACAA 120

Db 1243 ACCATCTCTCCACAA 1258

RESULT 11
US-09-067-615-123
Sequence 123, Application US/09067615
Patent No. 630645
GENERAL INFORMATION:
APPLICANT: Rhode, Peter R.
APPLICANT: Jiao, Jin-An
APPLICANT: Burkhardt, Martin
APPLICANT: Wong, Hing
TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dade International, Inc.
STREET: 1717 Deerfield Road
CITY: Deerfield
STATE: Illinois
COUNTRY: USA
ZIP: 60015
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/067,615
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/596,387
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/382,454
FILING DATE: 01-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,302
FILING DATE: 29-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pearson, Louise S.
REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STR-4665-CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 267-5300
TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 123:
SEQUENCE CHARACTERISTICS:
LENGTH: 1382 base pairs
TYPE: nucleic acid

STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 6..1382
US-09-067-615-123

Query Match 12.6%; Score 16; DB 4; Length 1382;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 ACCATTCTTCACAA 120
DB 1243 ACCATTCTTCACAA 1258

RESULT 12
PCT-US95-09816A-123
Sequence 123, Application PC/TUS9509816A
GENERAL INFORMATION:
APPLICANT: Wong, Hing C.
APPLICANT: Rhode, Peter R.
APPLICANT: Widanz, Jon A.
APPLICANT: Grammer, Susan
APPLICANT: Edwards, Ana C.
APPLICANT: Chavallaz, Pierre-Andre
APPLICANT: Jiao, Jin-An
TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dade International, Inc.
STREET: 1717 Deerfield Road
CITY: Deerfield
STATE: Illinois
COUNTRY: USA
ZIP: 60015
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09816A
FILING DATE: 31-JUL-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/382,454
FILING DATE: 01-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,302
FILING DATE: 29-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pearson, Louise S.
REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STR-4665-CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 267-5300
TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 123:
SEQUENCE CHARACTERISTICS:
LENGTH: 1382 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 6..1382
PCT-US95-09816A-123
Query Match 12.6%; Score 16; DB 5; Length 1382;
Best Local Similarity 100.0%; Pred. No. 20;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 105 ACCATTCTTCACAA 120
DB 1243 ACCATTCTTCACAA 1258

RESULT 13
US-08-596-387B-121
Sequence 121, Application US/08596387B
Patent No. 5869270
GENERAL INFORMATION:
APPLICANT: Rhode, Peter R.
APPLICANT: Jiao, Jin-An
APPLICANT: Burkhardt, Martin
APPLICANT: Wong, Hing
TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dade International, Inc.
STREET: 1717 Deerfield Road
CITY: Deerfield
STATE: Illinois
COUNTRY: USA
ZIP: 60015
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/596,387B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09816
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/382,454
FILING DATE: 01-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,302
FILING DATE: 29-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pearson, Louise S.
REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STR-4665-CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 267-5300
TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 1385 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 6..1382
US-08-596-387B-121
Query Match 12.6%; Score 16; DB 2; Length 1385;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 105 ACCATTCTTCACAA 120
DB 1243 ACCATTCTTCACAA 1258
RESULT 14
US-09-067-615-121

; Sequence 121, Application US/09067615
; Patent No. 6309645
; GENERAL INFORMATION:
; APPLICANT: Rhode, Peter R.
; APPLICANT: Jiao, Jin-An
; APPLICANT: Burkhardt, Martin
; APPLICANT: Wong, Hing
; TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dade International, Inc.
; STREET: 1717 Deerfield Road
; CITY: Deerfield
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60015
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/067,615
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/596,387
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/382,454
; FILING DATE: 01-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,302
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Pearson, Louise S.
; REGISTRATION NUMBER: 32,369
; REFERENCE/DOCKET NUMBER: STR-4665-CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 267-5376
; TELEFAX: (708) 267-5376
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..1382
; US-09-067-615-121

Query Match 12.6%; Score 16; DB 4; Length 1385;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 105 ACCATTCCTTCACAA 120
|||
Db 1243 ACCATTCCTTCACAA 1258

RESULT 15
PCT-US95-09816A-121
; Sequence 121, Application PC/TUS9509816A
; GENERAL INFORMATION:
; APPLICANT: Mong, Hing C.
; APPLICANT: Rhode, Peter R.
; APPLICANT: Widanz, Jon A.
; APPLICANT: Grammer, Susan
; APPLICANT: Edwards, Ana C.
; APPLICANT: Chavallaz, Pierre-Andre
; APPLICANT: Jiao, Jin-An

; TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dade International, Inc.
; STREET: 1717 Deerfield Road
; CITY: Deerfield
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60015
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09816A
; FILING DATE: 31-JUL-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/382,454
; FILING DATE: 01-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,302
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Pearson, Louise S.
; REGISTRATION NUMBER: 32,369
; REFERENCE/DOCKET NUMBER: STR-4665-CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 267-5376
; TELEFAX: (708) 267-5376
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..1382
; PCT-US95-09816A-121

Query Match 12.6%; Score 16; DB 5; Length 1385;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 105 ACCATTCCTTCACAA 120
|||
Db 1243 ACCATTCCTTCACAA 1258

RESULT 16
US-08-596-387B-122
; Sequence 122, Application US/08596387B
; Patent No. 5869270
; GENERAL INFORMATION:
; APPLICANT: Rhode, Peter R.
; APPLICANT: Jiao, Jin-An
; APPLICANT: Burkhardt, Martin
; APPLICANT: Wong, Hing
; TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dade International, Inc.
; STREET: 1717 Deerfield Road
; CITY: Deerfield
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60015
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/596,387B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09816
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/382,454
FILING DATE: 01-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,302
FILING DATE: 29-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pearson, Louise S.
REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STR-4665-CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 267-5300
TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 6..1505
US-08-596-387B-122

Query Match 12.6%; Score 16; DB 2; Length 1508;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 ACCATTCCTCCACAA 120
DB 1243 ACCATTCCTCCACAA 1258

RESULT 17
US-08-960-190A-24
Sequence 24, Application US/08960190A
Patent No. 6232445
GENERAL INFORMATION:
APPLICANT: Rhode, Peter R.
APPLICANT: Acevedo, Jorge
APPLICANT: Burkhardt, Martin
APPLICANT: Jiao, Jin-An
APPLICANT: Wong, Hing C.
TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND
METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: usa
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,190A
FILING DATE: 29-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cortless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 6..1505
OTHER INFORMATION:
US-08-960-190A-24

Query Match 12.6%; Score 16; DB 3; Length 1508;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 ACCATTCCTCCACAA 120
DB 1243 ACCATTCCTCCACAA 1258

RESULT 18
US-09-067-615-122
Sequence 122, Application US/09067615
Patent No. 6309645
GENERAL INFORMATION:
APPLICANT: Rhode, Peter R.
APPLICANT: Burkhardt, Martin
APPLICANT: Jiao, Jin-An
APPLICANT: Wong, Hing
TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dade International, Inc.
STREET: 1717 Deerfield Road
CITY: Deerfield
STATE: Illinois
COUNTRY: USA
ZIP: 60015
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/067,615
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/596,387
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/382,454
FILING DATE: 01-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,302
FILING DATE: 29-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pearson, Louise S.
REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STR-4665-CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 267-5300

TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 6..1505
US-09-067-615-122

Query Match 12.6%; Score 16; DB 4; Length 1508;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 ACCATTCCTTCACAA 120
DB 1243 ACCATTCCTTCACAA 1258

RESULT 19
PCT-US95-09816A-122
Sequence 122, Application PC/TUS9509816A
GENERAL INFORMATION:
APPLICANT: Wong, Hing C.
APPLICANT: Rhode, Peter R.
APPLICANT: Widanz, Jon A.
APPLICANT: Grammer, Susan
APPLICANT: Edwards, Ana C.
APPLICANT: Chavallaz, Pierre-Andre
APPLICANT: Jiao, Jin-An
TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dade International, Inc.
STREET: 1717 Deerfield Road
CITY: Deerfield
STATE: Illinois
COUNTRY: USA
ZIP: 60015
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09816A
FILING DATE: 31-JUL-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/382,454
FILING DATE: 01-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,302
FILING DATE: 29-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pearson, Louise S.
REGISTRATION NUMBER: 32,369
REFERENCE/DOCKET NUMBER: STP-4665-CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 267-5300
TELEFAX: (708) 267-5376
INFORMATION FOR SEQ ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS

LOCATION: 6..1505
PCT-US95-09816A-122

Query Match 12.6%; Score 16; DB 5; Length 1508;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 ACCATTCCTTCACAA 120
DB 1243 ACCATTCCTTCACAA 1258

RESULT 20
US-09-780-016-27
Sequence 27, Application US/09780016
Patent No. 6509456
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedlich, Glenn
APPLICANT: Abuin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6509456el Human Proteases and
FILE REFERENCE: LEX-0132-USA
CURRENT APPLICATION NUMBER: US/09/780,016
CURRENT FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,294
PRIOR FILING DATE: 2000-02-11
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 3208
TYPE: DNA
ORGANISM: homo sapiens
US-09-780-016-27

Query Match 12.6%; Score 16; DB 4; Length 3208;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CTGTGCATCTGCAGG 25
DB 2820 CTGTGCATCTGCAGG 2835

RESULT 21
US-09-194-285-7
Sequence 7, Application US/09194285
Patent No. 6355479
GENERAL INFORMATION:
APPLICANT: Webb, Susan R.
APPLICANT: Winovist, Ola
APPLICANT: Karlsson, Lars
APPLICANT: Jackson, Michael R.
TITLE OF INVENTION: MHC Class II Antigen Presenting Systems
FILE REFERENCE: TSRI 536.1
CURRENT APPLICATION NUMBER: US/09/194,285
CURRENT FILING DATE: 1999-04-12
PRIOR APPLICATION NUMBER: PCT/US97/08697
PRIOR FILING DATE: 1997-05-22
PRIOR APPLICATION NUMBER: US 60/018,175
PRIOR FILING DATE: 1996-05-23
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 4713
TYPE: DNA
ORGANISM: Mus musculus

US-09-194-285-7

Query Match 12.6%; Score 16; DB 4; Length 4713;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 ACCATTCTTCACAA 120

DB 964 ACCATTCTTCACAA 979

RESULT 22

US-08-652-971-1/c

; Sequence 1, Application us/08652971

; Patent No. 5814507

; GENERAL INFORMATION:

; APPLICANT: Cheng, Jill

; APPLICANT: Lasky, Laurence A.

; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE

; TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd.

; CITY: South San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/652,971

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Dregger, Ginger R.

; REGISTRATION NUMBER: 33,055

; REFERENCE/DOCKET NUMBER: P1033

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 225-3216

; TELEFAX: (415) 952-9881

; TELEX: 910 371-7168

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5769 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 379..4686

; US-08-652-971-1

Query Match 12.6%; Score 16; DB 1; Length 5769;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 TTGTGCTCGGGAGA 99

DB 890 TTGTGCTCGGGAGA 875

RESULT 23

US-08-991-258A-1/c

; Sequence 1, Application US/08991258A

; Patent No. 5928887

; GENERAL INFORMATION:

; APPLICANT: Cheng, Jill

; APPLICANT: Lasky, Laurence A.

; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE

; TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT, LLP

; STREET: 4 Embarradero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/991,258A

; FILING DATE: 17-DEC-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/652,971

; FILING DATE: 24-MAY-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Dregger, Walter H.

; REGISTRATION NUMBER: 24,190

; REFERENCE/DOCKET NUMBER: A-63478-3/WH/MTK

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5769 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 379..4686

; US-08-991-258A-1

Query Match 12.6%; Score 16; DB 2; Length 5769;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 TTGTGCTCGGGAGA 99

DB 890 TTGTGCTCGGGAGA 875

RESULT 24

US-08-769-399-1/c

; Sequence 1, Application US/08769399

; Patent No. 5976852

; GENERAL INFORMATION:

; APPLICANT: Cheng, Jill

; APPLICANT: Lasky, Laurence A.

; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE

; TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd.

; CITY: South San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/769,399
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Dreger, Ginger R.
;; REGISTRATION NUMBER: 33,055
;; REFERENCE/DOCKET NUMBER: P1033
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 225-3216
;; TELEFAX: (415) 952-9881
;; TELEEX: 910 371-7168
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5769 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 379..4686
;;
US-08-769-399-1

Query Match 12.6%; Score 16; DB 2; Length 5769;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 TTGTGCTGTGGGAGGA 99
|||
Db 890 TTGTGCTGTGGGAGGA 875

RESULT 25
US-08-991-953A-1/c
; Sequence 1, Application US/08991953A
; Patent No. 6083748
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT, LLP
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,953A
; FILING DATE: 16-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,971
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-63478-3/WHO/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5769 base pairs

;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 379..4686
;;
US-08-991-953A-1

Query Match 12.6%; Score 16; DB 3; Length 5769;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 TTGTGCTGTGGGAGGA 99
|||
Db 890 TTGTGCTGTGGGAGGA 875

RESULT 26
US-08-663-082-1
; Sequence 1, Application US/08663082
; Patent No. 6043411
; GENERAL INFORMATION:
; APPLICANT: NISHIZAWA, Osamu
; APPLICANT: TOGURI, Toshihiro
; TITLE OF INVENTION: GENE FOR FATTY ACID DESATURASE, VECTOR
; TITLE OF INVENTION: CONTAINING SAID GENE, PLANT TRANSFORMED WITH SAID GENE,
; TITLE OF INVENTION: AND PROCESS FOR CREATING SAID PLANT
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/663,082
; FILING DATE: 25-JUN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/02288
; FILING DATE: 28-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 93/352858
; FILING DATE: 28-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 81356/107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 196 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Anabaena variabilis
; STRAIN: IAM M-3
;;
US-08-663-082-1

Query Match 11.8%; Score 15; DB 3; Length 196;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 TTGTGCTCTGGGAG 98
|||||
Db 41 TTGTGCTCTGGGAG 55

RESULT 27
US-09-328-352-2677
; Sequence 2677, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2677
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2677

Query Match 11.8%; Score 15; DB 4; Length 219;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 CTGAAACACAGCAG 74
|||||
Db 101 CTGAAACACAGCAG 115

RESULT 28
US-09-313-294A-3641/C
; Sequence 3641, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 3641
; LENGTH: 266
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700612036H1
; NAME/KEY: unsure
; LOCATION: 128
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-3641

Query Match 11.8%; Score 15; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 GATGTCAGCTGAA 65
|||||
Db 154 GATGTCAGCTGAA 140

RESULT 29
US-09-702-705-1559
; Sequence 1559, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongrong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriek
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1559
; LENGTH: 266
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-1559

Query Match 11.8%; Score 15; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 59 GCTGAAACACAGCA 73
|||||
Db 196 GCTGAAACACAGCA 210

RESULT 30
US-09-736-457-1559.
; Sequence 1559, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongrong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriek
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1559
; LENGTH: 266
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-1559

Query Match 11.8%; Score 15; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 59 GCTGAAACACAGCA 73
|||||
Db 196 GCTGAAACACAGCA 210

RESULT 31
US-09-313-294A-2588/C
; Sequence 2588, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:

APPLICANT: Lalgudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO: 2588
LENGTH: 275
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700552942H1
US-09-313-294A-2588

Query Match 11.8%; Score 15; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAAGGAGCTGTGTT 85
DB 178 CAAGGAGCTGTGTT 164

RESULT 32
US-09-016-434-1031
Sequence 1031 Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1031:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTUT03
CLONE: 961450

US-09-016-434-1031

Query Match 11.8%; Score 15; DB 4; Length 318;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CTGTGCATCTGCAG 24
DB 205 CTGTGCATCTGCAG 219

RESULT 33
US-08-412-614-93
Sequence 93 Application US/08412614
Patent No. 553638
GENERAL INFORMATION:
APPLICANT: Rossau, Rudi
APPLICANT: Van Heuverswyn, Hugo
TITLE OF INVENTION: Hybridization Probes Derived from the
TITLE OF INVENTION: Spacer Region Between the 16S and 23S rRNA Genes for the
TITLE OF INVENTION: Detection of No. 553638-Viral Microorganisms
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 3100 No. 553638west Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402-4131
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch, 1.44 mb capacity
COMPUTER: IBM PC compatible (Compaq Deskpro 286e)
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect Version #5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/412,614
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,394
FILING DATE: 17-DEC-1992
APPLICATION NUMBER: PCT/EP91/00743
FILING DATE: 18-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB/90901054.3
FILING DATE: 18-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hillson, Randall A.
REGISTRATION NUMBER: 31,838
REFERENCE/DOCKET NUMBER: 8076.75-USWO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Haemophilus influenzae
STRAIN: NCTC 8143
US-08-412-614-93

Query Match 11.8%; Score 15; DB 1; Length 474;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 CAAGCTGAACACAG 70
|||||

Db 258 CAAGCTGAAAAACAAG 272

RESULT 34

US-08-635-761-93
Sequence 93, Application US/08635761
Patent No. 5945282

GENERAL INFORMATION:

APPLICANT: Rossau, Rudi
TITLE OF INVENTION: HYBRIDIZATION PROBES DERIVED FROM THE SPACER REGION BE
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 5945282west Center, 90 S. 7th Street
CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A.
ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/635,761
FILING DATE: 19-APR-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/965,394
FILING DATE: 17-DEC-1992
APPLICATION NUMBER: 08/412,614
FILING DATE: 29-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hillson, Randall A
REGISTRATION NUMBER: 31,838
REFERENCE/DOCKET NUMBER: 8076.75USC1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612/332-5300
TELEFAX: 612/332/9081
TELEX:

INFORMATION FOR SEQ ID NO: 93:

SEQUENCE CHARACTERISTICS:
LENGTH: 474 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-635-761-93

Query Match 11.8%; Score 15; DB 2; Length 474;

Best Local Similarity 100.0%; Pred. No. 68;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 CAAGCTGAAAAACAAG 70
|||
Db 258 CAAGCTGAAAAACAAG 272

RESULT 35

US-09-312-520-93
Sequence 93, Application US/09312520
Patent No. 6277577

GENERAL INFORMATION:

APPLICANT: Rossau, Rudi
TITLE OF INVENTION: HYBRIDIZATION PROBES DERIVED FROM THE SPACER REGION BE
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 6277577west Center, 90 S. 7th Street

CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A.
ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312,520
FILING DATE: 19-APR-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/965,394
FILING DATE: 17-DEC-1992
APPLICATION NUMBER: 08/412,614
FILING DATE: 29-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hillson, Randall A
REGISTRATION NUMBER: 31,838
REFERENCE/DOCKET NUMBER: 8076.75USC1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612/332-5300
TELEFAX: 612/332/9081
TELEX:

INFORMATION FOR SEQ ID NO: 93:

SEQUENCE CHARACTERISTICS:
LENGTH: 474 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-09-312-520-93

Query Match 11.8%; Score 15; DB 3; Length 474;

Best Local Similarity 100.0%; Pred. No. 68;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 CAAGCTGAAAAACAAG 70
|||
Db 258 CAAGCTGAAAAACAAG 272

RESULT 36

US-08-412-614-94
Sequence 94, Application US/08412614
Patent No. 5536638

GENERAL INFORMATION:

APPLICANT: Rossau, Rudi
TITLE OF INVENTION: Hybridization Probes Derived from the
TITLE OF INVENTION: Spacer Region Between the 16S and 23S rRNA Genes for the
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 3100 No. 5536638west Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402-4131

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 3.5 inch, 1.44 mb capacity
COMPUTER: IBM PC compatible (Compaq Deskpro 286e)
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect Version #5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/412,614

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,394
FILING DATE: 17-DEC-1992
APPLICATION NUMBER: PCT/EP91/00743
FILING DATE: 18-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB/90901054.3
FILING DATE: 18-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hillson, Randall A.
REGISTRATION NUMBER: 31,838
REFERENCE/DOCKET NUMBER: 8076.75-USWO.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Haemophilus influenzae
STRAIN: ITM 859
US-08-412-614-94

Query Match 11.8%; Score 15; DB 1, Length 476;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 CAAGCTGAAACAG 70
DB 258 CAAGCTGAAACAG 272

RESULT 37
US-08-635-761-94
Sequence 94, Application US/08635761
Patent No. 5945282
GENERAL INFORMATION:
APPLICANT: Rosseau, Rudi
TITLE OF INVENTION: HYBRIDIZATION PROBES DERIVED FROM THE SPACER REGION BE
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 5945282west Center, 90 S. 7th Street
CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A.
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/635,761
FILING DATE: 19-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,394
FILING DATE: 17-DEC-1992
APPLICATION NUMBER: 08/412,614
FILING DATE: 29-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hillson, Randall A.
REGISTRATION NUMBER: 31,838
REFERENCE/DOCKET NUMBER: 8076.75USC1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/332-5300
TELEFAX: 612/332/9081
TELEX:
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-635-761-94

Query Match 11.8%; Score 15; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 CAAGCTGAAACAG 70
DB 258 CAAGCTGAAACAG 272

RESULT 38
US-09-312-520-94
Sequence 94, Application US/09312520
Patent No. 6277577
GENERAL INFORMATION:
APPLICANT: Rosseau, Rudi
TITLE OF INVENTION: HYBRIDIZATION PROBES DERIVED FROM THE SPACER REGION BE
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 6277577west Center, 90 S. 7th Street
CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A.
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312,520
FILING DATE: 19-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,394
FILING DATE: 17-DEC-1992
APPLICATION NUMBER: 08/412,614
FILING DATE: 29-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hillson, Randall A.
REGISTRATION NUMBER: 31,838
REFERENCE/DOCKET NUMBER: 8076.75USC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/332-5300
TELEFAX: 612/332/9081
TELEX:
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:

ORIGINAL SOURCE:
US-09-312-520-94

Query Match 11.8%; Score 15; DB 3; Length 476;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 CAAGCTGAAACAG 70
|||||
DB 258 CAAGCTGAAACAG 272

RESULT 39

US-08-309-182B-2

Sequence 2, Application US/08309182B

Patent No. 5639645

GENERAL INFORMATION:

APPLICANT: No. 563964510 MURATA

TITLE OF INVENTION: A RECOMBINANT 9 DESATURASE AND A GENE

TITLE OF INVENTION: ENCODING THE SAME

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 500 kb

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/309,182B

FILING DATE: September 20, 1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek, Jr.

REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850

TELEFAX:

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 819 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

ORIGINAL SOURCE:

ORGANISM: Anabaena variabilis

FEATURE:

NAME/KEY: CDS

LOCATION: 1..816

IDENTIFICATION METHOD: P

US-08-309-182B-2

Query Match 11.8%; Score 15; DB 1; Length 819;

Best Local Similarity 100.0%; Pred. No. 68;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 TTGTGCTCTGGGAG 98
|||||
DB 521 TTGTGCTCTGGGAG 535

RESULT 40

US-09-121-425-3

Sequence 3, Application US/09121425

Patent No. 6153418

GENERAL INFORMATION:

APPLICANT: Lehmann, Martin

TITLE OF INVENTION: Consensus Phytases

FILE REFERENCE: consensus phytases 13239

CURRENT APPLICATION NUMBER: US/09/121,425

CURRENT FILING DATE: 1998-07-23

EARLIER APPLICATION NUMBER: EPO 97112688.3

EARLIER FILING DATE: 1997-07-24

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 1426

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:consensus

US-09-121-425-3

Query Match 11.8%; Score 15; DB 3; Length 1426;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 ATGTCAAGCTGAAA 66
|||||
DB 1250 ATGTCAAGCTGAAA 1264

RESULT 41

US-09-634-493A-3

Sequence 3, Application US/09634493A

Patent No. 6579975

GENERAL INFORMATION:

APPLICANT: Lehmann, Martin

TITLE OF INVENTION: Consensus Phytases

FILE REFERENCE: consensus phytases 13239

CURRENT APPLICATION NUMBER: US/09/634,493A

CURRENT FILING DATE: 2000-08-08

PRIOR APPLICATION NUMBER: US/09/121,425

PRIOR FILING DATE: 1998-07-23

PRIOR APPLICATION NUMBER: EPO 97112688.3

PRIOR FILING DATE: 1997-07-24

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 1426

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:consensus

US-09-634-493A-3

Query Match 11.8%; Score 15; DB 4; Length 1426;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 ATGTCAAGCTGAAA 66
|||||
DB 1250 ATGTCAAGCTGAAA 1264

RESULT 42

US-08-389-668A-1/c

Sequence 1, Application US/08389668A

Patent No. 5637470

GENERAL INFORMATION:

APPLICANT: Kaczorowski, Gregory J.

APPLICANT: Garcia, Maria L.

APPLICANT: Leonard, Reid J.

APPLICANT: McManus, Owen B.
APPLICANT: Swanson, Richard J.
APPLICANT: Folander, Kimberly L.
TITLE OF INVENTION: NOVEL HETEROMULTIMER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Francis P. Bigley
STREET: 126 E. Lincoln Avenue, P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/389,668A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/242,811
FILING DATE: 13-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bigley, Francis P.
REGISTRATION NUMBER: 36,356
REFERENCE/DOCKET NUMBER: 19226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-1249
TELEFAX: (908) 594-4270
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2238 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-389-668A-1

Query Match 11.8%; Score 15; DB 1; Length 2238;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 AAGCTGAAAACAGC 71
|||||
Db 1106 AAGCTGAAAACAGC 1092

RESULT 43
US-08-732-506-1/c
Sequence 1, Application US/08732506
Patent No. 5776734
GENERAL INFORMATION:
APPLICANT: Kaczorowski, Gregory J.
APPLICANT: Kaczorowski, Maria L.
APPLICANT: Leonard, Reid J.
APPLICANT: McManus, Owen B.
APPLICANT: Swanson, Richard J.
APPLICANT: Folander, Kimberly L.
TITLE OF INVENTION: NOVEL HETEROMULTIMER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Francis P. Bigley
STREET: 126 E. Lincoln Avenue, P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05768
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05768
FILING DATE:
FILING DATE: 13-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bigley, Francis P.
REGISTRATION NUMBER: 36,356
REFERENCE/DOCKET NUMBER: 19226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-1249
TELEFAX: (908) 594-4270
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2238 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-732-506-1

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/732,506
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05768
FILING DATE:
FILING DATE: 13-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bigley, Francis P.
REGISTRATION NUMBER: 36,356
REFERENCE/DOCKET NUMBER: 19226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-1249
TELEFAX: (908) 594-4270
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2238 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-732-506-1

Query Match 11.8%; Score 15; DB 1; Length 2238;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 AAGCTGAAAACAGC 71
|||||
Db 1106 AAGCTGAAAACAGC 1092

RESULT 44
PCT-US95-05768-1/c
Sequence 1, Application PC/TUS9505768
GENERAL INFORMATION:
APPLICANT: Kaczorowski, Gregory J.
APPLICANT: Kaczorowski, Maria L.
APPLICANT: Leonard, Reid J.
APPLICANT: McManus, Owen B.
APPLICANT: Swanson, Richard J.
APPLICANT: Folander, Kimberly L.
TITLE OF INVENTION: NOVEL HETEROMULTIMER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Francis P. Bigley
STREET: 126 E. Lincoln Avenue, P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05768
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/242,811
FILING DATE: 13-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bigley, Francis P.
REGISTRATION NUMBER: 36,356
REFERENCE/DOCKET NUMBER: 19226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-1249
TELEFAX: (908) 594-4270

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2238 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
PCT-US95-05768-1

Query Match 11.8%; Score 15; DB 5; Length 2238;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 AAGCTGAAAACAAGC 71
DB 1106 AAGCTGAAAACAAGC 1092

RESULT 45
US-07-959-943-10/c
Sequence 10, Application US/07959943
Patent No. 5418162
GENERAL INFORMATION:
APPLICANT: Blakey, Randy D.
APPLICANT: Fremieu Jr., Robert T.
APPLICANT: Caron, Marc G.
TITLE OF INVENTION: Serotonin Transporter cDNA
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth D. Sibley, Bell, Seltzer, Park
ADDRESS: and
ADDRESS: Gibson
STREET: Post Office Drawer 31107
CITY: Raleigh
STATE: No. 5418162ch Carolina
COUNTRY: U.S.A.
ZIP: 27622
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/959,943
FILING DATE: 19921014
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405.38a
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-881-3140
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2508 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
POSITION IN GENOME:
UNITS: 2278 basepairs
FEATURE:
NAME/KEY: CDS
LOCATION: 73..1962
US-07-959-943-10

Query Match 11.8%; Score 15; DB 1; Length 2508;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 23 AAGGTCAGGTGATG 37
|||||

DB 737 AAGGTCAGGTGATG 723

RESULT 46
US-09-074-658-69/c
Sequence 69, Application US/09074658
Patent No. 6184371
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M
APPLICANT: Run-Pan Du
APPLICANT: Quljun Wang
APPLICANT: Yang, Yan-Ping
APPLICANT: Klein, Michel H
TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/074,658
FILING DATE: 08-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-795
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 2718 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-074-658-69

Query Match 11.8%; Score 15; DB 3; Length 2718;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TACCTGTGCATCTG 21
DB 442 TACCTGTGCATCTG 428

RESULT 47
US-09-620-312D-552/c
Sequence 552, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundl, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunging

APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhwei
APPLICANT: John Tillinghaast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317.
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 552
LENGTH: 3582
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (392)..(2647)
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(3582)
OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-552

Query Match 11.8%; Score 15; DB 4; Length 3582;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 TGGGAGAGTGTAC 106
DB 3014 TGGGAGAGTGTAC 3000

RESULT 48
US-08-026-138E-17/c
Sequence 17, Application US/08026138E
Patent No. 5502166
GENERAL INFORMATION:
APPLICANT: Masayoshi MISHINA
TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nishiohata Residence 1-107
STREET: 5214, Nishiohata-machi
CITY: Niigata-shi
STATE: Niigata-ken
COUNTRY: JAPAN
ZIP: 951
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS V.5
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/026,138E
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 39563/1992
FILING DATE: 26-FEB-1992
APPLICATION NUMBER: JP 173155/1992
FILING DATE: 30-JUN-1992
APPLICATION NUMBER: JP 215017/1992
FILING DATE: 12-AUG-1992
APPLICATION NUMBER: JP 303878/1992
FILING DATE: 13-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hamburg, C. Bruce
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-4551

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 986-2340
TELEFAX: (212) 953-7733
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 4368 nucleic acids
TYPE: nucleic acid
STRANDEDNESS: double strand
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: mouse
TISSUE TYPE: brain
PUBLICATION INFORMATION:
AUTHORS: Masayoshi MISHINA
TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
RELEVANT RESIDUES IN SEQ ID NO: 17: FROM 1 to 4368
US-08-026-138E-17

Query Match 11.8%; Score 15; DB 1; Length 4368;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 TTGTGCTGGGAG 98
DB 3281 TTGTGCTGGGAG 3267

RESULT 49
US-08-026-138E-6/c
Sequence 6, Application US/08026138E
Patent No. 5502166
GENERAL INFORMATION:
APPLICANT: Masayoshi MISHINA
TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nishiohata Residence 1-107
STREET: 5214, Nishiohata-machi
CITY: Niigata-shi
STATE: Niigata-ken
COUNTRY: JAPAN
ZIP: 951
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS V.5
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/026,138E
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 39563/1992
FILING DATE: 26-FEB-1992
APPLICATION NUMBER: JP 173155/1992
FILING DATE: 30-JUN-1992
APPLICATION NUMBER: JP 215017/1992
FILING DATE: 12-AUG-1992
APPLICATION NUMBER: JP 303878/1992
FILING DATE: 13-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hamburg, C. Bruce
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-4551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 986-2340
TELEFAX: (212) 953-7733
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 4446 nucleic acids
TYPE: nucleic acid
STRANDEDNESS: double strand
TOPOLOGY: linear

MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: mouse
TISSUE TYPE: brain
PUBLICATION INFORMATION:
AUTHORS: Maayoshi MISHINA
TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 to 4446
US-08-026-138E-6

Query Match 11.8%; Score 15; DB 1; Length 4446;
Best Local Similarity 100.0%; Pred. No. 70;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 TTGTGCTGTGGGAG 98
|||||
Db 3359 TTGTGCTGTGGGAG 3345

RESULT 50

US-09-215-694-20/c

Sequence 20. Application US/09215694B

Patent No. 6391583

GENERAL INFORMATION:

APPLICANT: Wisconsin Alumni Research Foundation

APPLICANT: Hutchinson, Charles R.

APPLICANT: Kennedy, Jonathan n.m.i

APPLICANT: Park, Cheonsaek n.m.i

TITLE OF INVENTION: METHOD OF PRODUCING ANTIHYPERCHOLESTEROLEMIC AGENTS

FILE REFERENCE: 960296.95718

CURRENT APPLICATION NUMBER: US/09/215,694B

CURRENT FILING DATE: 1999-12-18

NUMBER OF SEQ ID NOS: 36

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 20

LENGTH: 5053

TYPE: DNA

ORGANISM: Aspergillus terreus

US-09-215-694-20

Query Match

Best Local Similarity

Matches 15; Conservative

OY 28 CCAGTGATGATGC 42

Db 210 CCAGTGATGATGC 196

Search completed: November 7, 2003, 11:59:33
Job time : 37.4064 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 11:58:22 ; Search time 489,785 Seconds
(without alignments)
827.407 Million cell updates/sec

Title: US-09-509-779-1_COPY_154_280

Perfect score: 127
Sequence: 1 GTCGATACCTGTGCATCT.....ATTCCCTCACAACCTGCTGC 127

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2141354 seqs, 1595478879 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	25.2	342	9	US-09-826-312-7
2	32	25.2	342	14	US-10-108-767-7
3	32	25.2	342	14	US-10-152-156-7
4	32	25.2	612	10	US-09-796-692-8547
5	32	25.2	612	14	US-10-040-863-8547
6	32	25.2	836	10	US-09-764-864-38
7	32	25.2	836	10	US-09-764-864-48
8	29	22.8	535	14	US-10-102-524-396
9	29	22.8	1024	14	US-10-198-846-12407
10	25	19.7	60	12	US-09-908-975-13285
11	25	16.5	933	14	US-10-198-846-7306
12	19	15.0	498	10	US-09-764-877-3347
13	19	15.0	3372	12	US-10-027-632-113318
14	19	15.0	3372	13	US-10-027-632-113318
15	18	14.2	441	12	US-10-027-632-138349
16	18	14.2	441	13	US-10-027-632-138349

C 17	18	14.2	833	12	US-10-027-632-162982	Sequence 162982.
C 18	18	14.2	833	12	US-10-027-632-162983	Sequence 162983.
C 19	18	14.2	833	13	US-10-027-632-162982	Sequence 162982.
C 20	18	14.2	833	13	US-10-027-632-162983	Sequence 162983.
C 21	18	14.2	930	12	US-10-027-632-31061	Sequence 31061, A
C 22	18	14.2	930	13	US-10-027-632-31061	Sequence 31061, A
C 23	17	13.4	252	12	US-10-027-632-276727	Sequence 276727.
C 24	17	13.4	252	13	US-10-027-632-276727	Sequence 276727.
C 25	17	13.4	348	10	US-09-878-574-2599	Sequence 2599, Ap
C 26	17	13.4	355	9	US-09-864-761-4974	Sequence 4974, Ap
C 27	17	13.4	355	9	US-09-864-761-21703	Sequence 21703, A
C 28	17	13.4	464	14	US-10-066-543-1656	Sequence 1656, Ap
C 29	17	13.4	481	10	US-09-878-574-1344	Sequence 1344, Ap
C 30	17	13.4	500	11	US-09-918-995-11855	Sequence 11855, A
C 31	17	13.4	575	12	US-10-027-632-224793	Sequence 224793.
C 32	17	13.4	575	13	US-10-027-632-224793	Sequence 224793.
C 33	17	13.4	712	12	US-10-027-632-145248	Sequence 145248.
C 34	17	13.4	712	13	US-10-027-632-145248	Sequence 145248.
C 35	17	13.4	1781	12	US-10-027-632-255720	Sequence 255720.
C 36	17	13.4	1781	13	US-10-027-632-255720	Sequence 255720.
C 37	17	13.4	2904	14	US-10-281-904-3	Sequence 3, Appli
C 38	17	13.4	3494	10	US-09-880-107-2295	Sequence 2295, Ap
C 39	17	13.4	3494	14	US-10-265-823-31	Sequence 31, Appl
C 40	17	13.4	3494	14	US-10-264-374-200	Sequence 200, App
C 41	17	13.4	3494	14	US-10-281-904-1	Sequence 1, Appli
C 42	17	13.4	3681	12	US-10-133-013-221	Sequence 221, App
C 43	17	13.4	3686	10	US-09-981-353-121	Sequence 121, App
C 44	17	13.4	3686	13	US-10-044-090-218	Sequence 218, App
C 45	17	13.4	3686	14	US-10-084-817-86	Sequence 86, Appl

ALIGNMENTS

RESULT 1
US-09-826-312-7
; Sequence 7, Application US/09826312
; Patent No. US20030042083A1
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jiating
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; TITLE OF INVENTION: UBIQUITIN LIGASE ASSAY
; FILE REFERENCE: A-68613-1/RMS/JTD
; CURRENT APPLICATION NUMBER: US/09/826.312
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 09/542.497
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-826-312-7

Query Match 25.2%; Score 32; DB 9; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 GCCATCTGCAGGTCACGATGATGATGCTG 45
DB 151 GCCATCTGCAGGTCACGATGATGATGCTG 182
RESULT 2
US-10-108-767-7
; Sequence 7, Application US/10108767
; Publication No. US20030104474A1
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jiating


```

; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8547
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (485)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-8547
```

```

Query Match      25.2%; Score 32; DB 14; Length 612;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      14 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 45
Db      35 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 66
```

```

RESULT 6
US-09-764-864-39
; Sequence 39, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-39
```

```

Query Match      25.2%; Score 32; DB 10; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      14 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 45
```

```

Db      166 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 197
```

```

RESULT 7
US-09-764-864-498
; Sequence 498, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 498
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (45)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-498
```

```

Query Match      25.2%; Score 32; DB 10; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      14 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 45
Db      219 GCCATCTGCAGGGTCCAGGTGATGATGCCTG 250
```

```

RESULT 8
US-10-102-524-396
; Sequence 396, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 396
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 14, 24, 32, 206, 240, 361, 384, 439, 442
; OTHER INFORMATION: n = A,T,C or G
US-10-102-524-396
```

```

Query Match      22.8%; Score 29; DB 14; Length 535;
Best Local Similarity 100.0%; Pred. No. 7.1e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      71 CAAGAGACTGTGTGTGTTCTTGGGAGA 99
Db      71 CAAGAGACTGTGTGTGTTCTTGGGAGA 99
```

```
RESULT 9
```

```
US-10-198-846-12407
; Sequence 12407, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12407
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 1022, 1023, 1024
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12407

Query Match
Best Local Similarity 22.8%; Score 29; DB 14; Length 1024;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAAGAGACTGTGTGTGTGTGTGGGAGA 99
DB 348 CAAGAGACTGTGTGTGTGTGTGGGAGA 376

RESULT 10
US-09-908-975-13285
; Sequence 13285, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: WASSERMAN, Avi
; APPLICANT: SHOSHAN, Avi
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13285
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-13285

Query Match
Best Local Similarity 19.7%; Score 25; DB 12; Length 60;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GCCATCTGCAGGCTCCAGGTGATGG 38
DB 36 GCCATCTGCAGGCTCCAGGTGATGG 60
```

```
RESULT 11
US-10-198-846-7306
; Sequence 7306, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7306
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 36, 150, 168, 187, 224, 266, 305, 375, 378, 405, 429, 438,
; LOCATION: 439, 454, 458, 464, 472, 485, 491, 495, 516, 536, 537, 538,
; LOCATION: 544, 545, 548, 550, 558, 559, 563, 565, 567, 570, 578, 581,
; LOCATION: 585, 617, 633, 635, 636, 637, 641, 644, 647, 651, 657
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 672, 674, 680, 690, 692, 694, 695, 699, 704, 705, 707, 709,
; LOCATION: 725, 733, 740, 743, 744, 747, 748, 754, 756, 757, 762, 765,
; LOCATION: 769, 771, 774, 775, 777, 783, 784, 786, 788, 792, 795, 797,
; LOCATION: 798, 801, 802, 805, 809, 812, 813, 814, 815, 817, 818
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 819, 824, 827, 836, 838, 846, 851, 852, 855, 860, 865, 871,
; LOCATION: 876, 882, 893, 901, 903, 904, 905, 908, 909, 910, 913, 915,
; LOCATION: 919, 925, 927, 930, 931
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-7306

Query Match
Best Local Similarity 16.5%; Score 21; DB 14; Length 933;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 CATTCTTCACACACTGCTGC 127
DB 93 CATTCTTCACACACTGCTGC 113

RESULT 12
US-09-764-877-3347
; Sequence 3347, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3347
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3347
```

Query Match 15.0%; Score 19; DB 10; Length 498;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 AACCAAGCAAGAGGACTGT 82
DB 403 AACCAAGCAAGAGGACTGT 421

RESULT 13

US-10-027-632-113318/c
; Sequence 113318, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113318
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113318

Query Match 15.0%; Score 19; DB 12; Length 3372;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 AGAGACTGTGTGTGTC 91
DB 1339 AGAGACTGTGTGTGTC 1321

RESULT 14

US-10-027-632-113318/c
; Sequence 113318, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113318
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113318

Query Match 15.0%; Score 19; DB 13; Length 3372;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 AGAGACTGTGTGTGTC 91
DB 1339 AGAGACTGTGTGTGTC 1321

RESULT 15

US-10-027-632-118349
; Sequence 118349, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118349
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-118349

Query Match 14.2%; Score 18; DB 12; Length 441;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 TCTGGGAGAGGTGTAAAC 107
DB 3 TCTGGGAGAGGTGTAAAC 20

RESULT 16

US-10-027-632-118349
; Sequence 118349, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

```

; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138349
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-138349

Query Match          14.2%: Score 18; DB 13; Length 441;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 TCTGGGAGAGGTGTAC 107
DB 3 TCTGGGAGAGGTGTAC 20

RESULT 17
US-10-027-632-162982
; Sequence 162982, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162982
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162982

Query Match          14.2%: Score 18; DB 12; Length 833;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 TGTCAAGCTGAAACAG 70
DB 194 TGTCAAGCTGAAACAG 211

RESULT 18
US-10-027-632-162983
; Sequence 162983, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162983
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162983

Query Match          14.2%: Score 18; DB 13; Length 833;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 TGTCAAGCTGAAACAG 70
DB 194 TGTCAAGCTGAAACAG 211

RESULT 19
US-10-027-632-162982
; Sequence 162982, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162982
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162982

Query Match          14.2%: Score 18; DB 13; Length 833;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 TGTCAAGCTGAAACAG 70
DB 194 TGTCAAGCTGAAACAG 211
```

Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 TGTCAAGCTGAACAAG 70
DB 194 TGTCAAGCTGAACAAG 211

RESULT 20

US-10-027-632-162983
Sequence 162983, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 162983
LENGTH: 833
TYPE: DNA
ORGANISM: Human
US-10-027-632-162983

Query Match 14.2%; Score 18; DB 13; Length 833;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 TGTCAAGCTGAACAAG 70
DB 194 TGTCAAGCTGAACAAG 211

RESULT 21

US-10-027-632-31061
Sequence 31061, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31061
LENGTH: 930
TYPE: DNA
ORGANISM: Human
US-10-027-632-31061

Query Match 14.2%; Score 18; DB 12; Length 930;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 ATTCTTCACACACTGCT 125
DB 774 ATTCTTCACACACTGCT 791

RESULT 22

US-10-027-632-31061
Sequence 31061, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31061
LENGTH: 930
TYPE: DNA
ORGANISM: Human
US-10-027-632-31061

Query Match 14.2%; Score 18; DB 13; Length 930;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 ATTCTTCACACACTGCT 125
DB 774 ATTCTTCACACACTGCT 791

RESULT 23

US-10-027-632-276727/c
Sequence 276727, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676


```

; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 276727
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-276727
```

```
Query Match          13.4% Score 17; DB 12; Length 252;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      21 GCAGGTCGAGTGATG 37
      |||||
DB      51 GCAGGTCGAGTGATG 35
```

```
RESULT 24
US-10-027-632-276727/c
; Sequence 276727, Application US/10027632
```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 276727
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-276727
```

```
Query Match          13.4% Score 17; DB 13; Length 252;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      21 GCAGGTCGAGTGATG 37
      |||||
DB      51 GCAGGTCGAGTGATG 35
```

```
RESULT 25
US-09-878-574-2599/c
; Sequence 2599, Application US/09878574
; Patent No. US20020110548A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO: 2559
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-023-Q1-B1-B1
US-09-878-574-2599
```

```
Query Match          13.4% Score 17; DB 10; Length 348;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      101 TGTAACTTCCTTCCA 117
      |||||
DB      205 TGTAACTTCCTTCCA 189
```

```
RESULT 26
US-09-864-761-4974
; Sequence 4974, Application US/09864761
```

```

; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
```

```
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 4974
/ LENGTH: 355
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AL031056.1
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
US-09-864-761-4974

Query Match      13.4%; Score 17; DB 9; Length 355;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      25 GGTCCAGGTGATGATG 41
Db      281 GGTCCAGGTGATGATG 297

RESULT 27
US-09-864-761-21703
/ Sequence 21703, Application US/09864761
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
/ FILE REFERENCE: Aecm1ca-X-1
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
```

```
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 21703
/ LENGTH: 355
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AL031056.1
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
/ OTHER INFORMATION: NT HIT: AB019786.1, EVALUATE 1.20e-02
/ OTHER INFORMATION: EST_HUMAN HIT: AA495851.1, EVALUATE 2.20e-01
US-09-864-761-21703

Query Match      13.4%; Score 17; DB 9; Length 355;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      25 GGTCCAGGTGATGATG 41
Db      281 GGTCCAGGTGATGATG 297

RESULT 28
US-10-066-543-1656
/ Sequence 1656, Application US/10066543
/ Publication No. US20030087818A1
/ GENERAL INFORMATION:
/ APPLICANT: Jiang, Yugu
/ APPLICANT: Pyle, Ruth A.
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Indirias, Carol Joseph
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Secrist, Heather
/ APPLICANT: Carter, Darrick
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Smith, Carole L.
/ APPLICANT: Durham, Margarita
/ APPLICANT: Stolk, John A.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
/ FILE REFERENCE: 210121.563
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR FILING DATE: 2002-01-31
/ NUMBER OF SEQ ID NOS: 3417
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1656
/ LENGTH: 464
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 211, 453, 454
/ OTHER INFORMATION: n = A,T,C or G
```

US-10-066-543-1656

Query Match 13.4%; Score 17; DB 14; Length 464;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CCATCTGAGGGTCCAG 31
|||||
DB 433 CCATCTGAGGGTCCAG 449

RESULT 29
US-09-878-574-1344/C

; Sequence 1344; Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 1344
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIR3028-039-Q1-B1-E10
US-09-878-574-1344

Query Match 13.4%; Score 17; DB 10; Length 481;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 TGTACCATTCCTTCCA 117
|||||
DB 439 TGTACCATTCCTTCCA 423

RESULT 30
US-09-918-995-31855/C

; Sequence 31855; Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31855
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(500)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31855

Query Match 13.4%; Score 17; DB 11; Length 500;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CCATCTGAGGGTCCAG 31

DB 482 CCATCTGAGGGTCCAG 466

RESULT 31
US-10-027-632-224793

; Sequence 224793; Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224793
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-224793

Query Match 13.4%; Score 17; DB 12; Length 575;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 CAAGCTGAACACAGCA 72
|||||
DB 46 CAAGCTGAACACAGCA 62

RESULT 32
US-10-027-632-224793

; Sequence 224793; Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224793

LENGTH: 575
TYPE: DNA
ORGANISM: Human
US-10-027-632-224793

Query Match 13.4%; Score 17; DB 13; Length 575;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 CAAGCTGAAACAGCA 72
DB 46 CAAGCTGAAACAGCA 62

RESULT 33

US-10-027-632-145248/C
Sequence 145248, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 145248

LENGTH: 712

TYPE: DNA

ORGANISM: Human

US-10-027-632-145248

Query Match 13.4%; Score 17; DB 12; Length 712;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ATCTGAGGTCACAGT 33
DB 689 ATCTGAGGTCACAGT 673

RESULT 34

US-10-027-632-145248/C
Sequence 145248, Application US/10027632
GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 145248

LENGTH: 712

TYPE: DNA

ORGANISM: Human

US-10-027-632-145248

Query Match 13.4%; Score 17; DB 13; Length 712;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ATCTGAGGTCACAGT 33
DB 689 ATCTGAGGTCACAGT 673

RESULT 35

US-10-027-632-255720
Sequence 255720, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 255720

LENGTH: 1781

TYPE: DNA

ORGANISM: Human

US-10-027-632-255720

Query Match 13.4%; Score 17; DB 12; Length 1781;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 TTCCTCCACACAGTCT 125
DB 1695 TTCCTCCACACAGTCT 1711

RESULT 36

US-10-027-632-255720
Sequence 255720, Application US/10027632
GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome

```
FILE REFERENCE: 108827,129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 255720
LENGTH: 1781
TYPE: DNA
ORGANISM: Human
US-10-027-632-255720
```

```
Query Match 13.4% Score 17; DB 13; Length 1781;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 109 TTCCTTCCACACTGCT 125
Db 1695 TTCCTTCCACACTGCT 1711
```

```
RESULT 37
US-10-281-904-3/c
Sequence 3, Application US/10281904
Publication No. US20030119036A1
GENERAL INFORMATION:
APPLICANT: Chun, Miyoung
TITLE OF INVENTION: METHODS OF USING 48149, A HUMAN
TITLE OF INVENTION: AMINOPEPTIDASE FAMILY MEMBER
FILE REFERENCE: MP101-174P1RM
CURRENT APPLICATION NUMBER: US/10/281,904
CURRENT FILING DATE: 2002-10-28
PRIOR APPLICATION NUMBER: 60/335,084
PRIOR FILING DATE: 2001-10-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 3
LENGTH: 2904
TYPE: DNA
ORGANISM: Homo sapien
US-10-281-904-3
```

```
Query Match 13.4% Score 17; DB 14; Length 2904;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 15 CCATCTGCAGGCTCCAG 31
Db 1642 CCATCTGCAGGCTCCAG 1626
```

```
RESULT 38
US-09-880-107-2295/c
Sequence 2295, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
```

```
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 2295
LENGTH: 3494
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 M22324
US-09-880-107-2295
```

```
Query Match 13.4% Score 17; DB 10; Length 3494;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 15 CCATCTGCAGGCTCCAG 31
Db 1762 CCATCTGCAGGCTCCAG 1746
```

```
RESULT 39
US-10-205-823-31/c
Sequence 31, Application US/10205823
Publication No. US20030108963A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarapu, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Monsey, Angela M.
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 31
LENGTH: 3494
TYPE: DNA
ORGANISM: Homo sapiens
US-10-205-823-31
```

```
Query Match 13.4% Score 17; DB 14; Length 3494;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 15 CCATCTGCAGGCTCCAG 31
Db 1762 CCATCTGCAGGCTCCAG 1746
```

Db 1762 CCATCTGCAGGGTCCAG 1746

RESULT 40
US-10-264-374-200/c
; Sequence 200, Application US/10264374
; Publication No. US20030113320A1
; GENERAL INFORMATION:
; APPLICANT: Rusolanti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-LJ 3203
; CURRENT APPLICATION NUMBER: US/10/264,374
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US/09/139,802
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 08/926,914
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 08/710,067
; PRIOR FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 200
; LENGTH: 3494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(3024)
US-10-264-374-200

Query Match 13.4%; Score 17; DB 14; Length 3494;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CCATCTGCAGGGTCCAG 31
Db 1762 CCATCTGCAGGGTCCAG 1746

RESULT 41
US-10-281-904-1/c
; Sequence 1, Application US/10281904
; Publication No. US20030119036A1
; GENERAL INFORMATION:
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: METHODS OF USING 48149, A HUMAN
; TITLE OF INVENTION: AMINOPEPTIDASE FAMILY MEMBER
; FILE REFERENCE: MP101-174P1RM
; CURRENT APPLICATION NUMBER: US/10/281,904
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: 60/335,084
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3494
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-281-904-1

Query Match 13.4%; Score 17; DB 14; Length 3494;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CCATCTGCAGGGTCCAG 31
Db 1762 CCATCTGCAGGGTCCAG 1746

RESULT 42

US-10-133-013-221/c
; Sequence 221, Application US/10133013
; Publication No. US20030166903A1
; GENERAL INFORMATION:
; APPLICANT: Astromoff, Anna
; APPLICANT: Bandman, Olga
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE
; FILE REFERENCE: PA-0049 US
; CURRENT APPLICATION NUMBER: US/10/133,013
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/287,067
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PERL Program
; SEQ ID NO 221
; LENGTH: 3681
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030166903A1 1136923.19
US-10-133-013-221

Query Match 13.4%; Score 17; DB 12; Length 3681;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CCATCTGCAGGGTCCAG 31
Db 1960 CCATCTGCAGGGTCCAG 1944

RESULT 43
US-09-981-353-121/c
; Sequence 121, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 121
; LENGTH: 3686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1344279CB1
US-09-981-353-121

Query Match 13.4%; Score 17; DB 10; Length 3686;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CCATCTGCAGGGTCCAG 31
Db 1962 CCATCTGCAGGGTCCAG 1946

RESULT 44
US-10-044-090-218/c
; Sequence 218, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090

```

; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 218
; LENGTH: 3686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1344279CB1
US-10-044-090-218
```

```

Query Match          13.4%; Score 17; DB 13; Length 3686;
Best Local Similarity 100.0%; Pred.No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      15 CCATCTGCAGGGTCCAG 31
DB      1962 CCATCTGCAGGGTCCAG 1946
```

```

RESULT 45
US-10-084-817-86/c
; Sequence 86, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Pion
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 86
; LENGTH: 3686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1344279CB1
US-10-084-817-86
```

```

Query Match          13.4%; Score 17; DB 14; Length 3686;
Best Local Similarity 100.0%; Pred.No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      15 CCATCTGCAGGGTCCAG 31
DB      1962 CCATCTGCAGGGTCCAG 1946
```

```

RESULT 46
US-09-925-297-338/c
; Sequence 338, Application US/09925297
; Patent No. US2002081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 338
```

```

; LENGTH: 3728
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3707)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3713)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-338
```

```

Query Match          13.4%; Score 17; DB 9; Length 3728;
Best Local Similarity 100.0%; Pred.No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      15 CCATCTGCAGGGTCCAG 31
DB      1972 CCATCTGCAGGGTCCAG 1956
```

```

RESULT 47
US-10-133-013-222/c
; Sequence 222, Application US/10133013
; Publication No. US20030166903A1
; GENERAL INFORMATION:
; APPLICANT: Astromoff, Anna
; APPLICANT: Bandman, Olga
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE
; FILE REFERENCE: PA-0049 US
; CURRENT APPLICATION NUMBER: US/10/133,013
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/287,067
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PERL Program
; SEQ ID NO 222
; LENGTH: 3815
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030166903A1 1136923.18
US-10-133-013-222
```

```

Query Match          13.4%; Score 17; DB 12; Length 3815;
Best Local Similarity 100.0%; Pred.No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      15 CCATCTGCAGGGTCCAG 31
DB      1961 CCATCTGCAGGGTCCAG 1945
```

```

RESULT 48
US-09-764-891-6245/c
; Sequence 6245, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAMM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6245
; LENGTH: 11150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-6245
```

Query Match 13.4%; Score 17; DB 11; Length 11150;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CCATCTGACGGGTCCAG 31
|||||
DB 76 CCATCTGACGGGTCCAG 60

RESULT 49

US-10-205-561/c
; Sequence 561, Application US/10205428
; Publication No. US20030108907A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P117C1
; CURRENT APPLICATION NUMBER: US/10/205,428
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1019
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 561
; LENGTH: 11150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-428-561

Query Match 13.4%; Score 17; DB 14; Length 11150;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CCATCTGACGGGTCCAG 31
|||||
DB 76 CCATCTGACGGGTCCAG 60

RESULT 50

US-09-795-668-1
; Sequence 1, Application US/09795668
; Patent No. US20020045577A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinhorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503841

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: r=g or a
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: y=c/u or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: m=a or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: k=g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s=g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: w=a or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: b=g or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: d=a or g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: h=a or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: v=a or g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: n=a or g or c or t/u
US-09-795-668-1

Query Match 13.4%; Score 17; DB 9; Length 1503841;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 CTGAAACACGACAGAG 76
|||||
DB 1166621 CTGAAACACGACAGAG 1166637

Search completed: November 7, 2003, 13:55:39
Job time : 491.785 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 02:14:16 ; Search time 1294.79 Seconds
(without alignments)
2383.909 Million cell updates/sec

Title: US-09-509-779-1_COPY_154_280

Perfect score: 127
Sequence: 1 GGGCATTACCTGTGCCATCT.....ATTCTTCACAACTGCTGC 127

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estcov:*
5: em_estcov:*
6: em_estcov:*
7: em_estcov:*
8: em_estcov:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_esthum:*
18: em_esthum:*
19: em_esthum:*
20: em_esthum:*
21: em_esthum:*
22: em_esthum:*
23: em_esthum:*
24: em_esthum:*
25: em_esthum:*
26: em_esthum:*
27: em_esthum:*
28: gb_est1:*
29: gb_est2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	100.0	263	10	BE624166 un42d01.y
2	127	100.0	291	14	CB570887 AGENCOURT
3	127	100.0	296	9	AA231201 mw38911.x
4	127	100.0	303	10	BE624155 un42c01.y

5	127	100.0	309	9	AA00941 mg36b06.x
6	127	100.0	310	13	BY343528
7	127	100.0	319	13	BY329058
8	127	100.0	321	14	W78263
9	127	100.0	322	13	BY125373
10	127	100.0	323	10	BE692680
11	127	100.0	325	10	BF012987
12	127	100.0	328	9	AA623723
13	127	100.0	328	13	BY342015
14	127	100.0	329	13	BY114710
15	127	100.0	332	9	AA110259
16	127	100.0	346	13	BY337380
17	127	100.0	348	13	BY106705
18	127	100.0	349	13	BY116440
19	127	100.0	351	13	BY170257
20	127	100.0	352	13	BY091208
21	127	100.0	355	10	BE089348
22	127	100.0	357	13	BY090965
23	127	100.0	361	13	BY302366
24	127	100.0	366	9	AA403365
25	127	100.0	367	9	AA144182
26	127	100.0	368	13	BY089817
27	127	100.0	371	13	BY091920
28	127	100.0	371	13	BY169149
29	127	100.0	376	13	BY172672
30	127	100.0	382	13	BY161786
31	127	100.0	392	13	BY307739
32	127	100.0	393	14	CD565301
33	127	100.0	394	13	BY278099
34	127	100.0	395	13	BY095585
35	127	100.0	397	10	BE277656
36	127	100.0	401	10	BE011540
37	127	100.0	409	14	W34101
38	127	100.0	414	9	AA499298
39	127	100.0	416	10	BE469845
40	127	100.0	418	13	BY238589
41	127	100.0	427	10	BE333776
42	127	100.0	429	10	BE692461
43	127	100.0	440	9	AA815479
44	127	100.0	443	9	AA239898
45	127	100.0	445	14	W34374

ALIGNMENTS

RESULT 1
BE624166
LOCUS
DEFINITION
un42d01.y1 Scores_mammary_gland_NMLMG Mus musculus cDNA clone IMAGE:3374593 5' similar to WP:R10A10.2 CE12670 ZINC FINGER PROTEIN
, mRNA sequence.

ACCESSION
BE624166
VERSION
KEYWORDS
SOURCE
ORGANISM

Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS
TITLE
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CCGAP),
Tumor Gene Index

JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
WGI:1084197

FEATURES
source
Seq primer: -40RP from Gibco
High quality sequence stop: 110.
Location/Qualifiers
1..263

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3374593"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_id="Soares-mammary_gland_NMLMG"
/notes="vector: pTVT3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pTVT3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaído."

```

Query Match	100.0%	Score 127	DB 10	Length 263
Best Local Similarity	100.0%	Pred. No. 1.7e-55		
Matches 127	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	GTGCATACCTGTGCCATCTGCAGGGATCCAGGTGATGTGATGCGCTGCTCATGTCTCAAGC	60	
Db	62	GTGGCATACCTGTGCCATCTGCAGGGATCCAGGTGATGTGATGCGCTGCTCATGTCTCAAGC	121	
Qy	61	TGAAAACAGCAAGAGACTGTGTGTGTGTGTGGGAGAGGTAAACAATTCCTTCACAA	120	
Db	122	TGAAAACAGCAAGAGACTGTGTGTGTGTGTGGGAGAGGTAAACAATTCCTTCACAA	181	
Qy	121	CTGCTGC	127	
Db	182	CTGCTGC	188	

RESULT 2									
CB570887									
LOCUS									
DEFINITION	CB570887		291 bp	mRNA	linear	EST 02-APR-2003			
ACCESSION	AGNCOCURT_13009672		NIH MGC_165 Mus musculus cDNA clone						
VERSION	IMAGE:30276894.5,		mrna sequence.						
KEYWORDS	CB570887								
SOURCE	CB570887.1		GI:29490417						
ORGANISM	EST.								
	Mus musculus (house mouse)								
	Mus musculus								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
AUTHORS	1 (bases 1 to 291)								
TITLE	NIH-MGC http://mgc.nci.nih.gov/.								
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)								
COMMENT	Unpublished								
	Contact: Robert Strausberg, Ph.D.								

FEATURES
SOURCE

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone=IMAGE:30276894"
/tissue_type="primary cultures of Sertoli cells"
/lab_host="MDH10 (T1-phage-resistant)"
/clone_lib="NH MGC 165"
/notes="Organ: testis; Vector: pDNR-LIB; Site_1: SfiI"

```

(ggcgcattatggcgc) ; Site 2: 5'fl (ggcgcgcctcggcc) ; 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCATTTATGAGC-Cr(3)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.6-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library.

Query Match	100.0%	Score 127	DB 14	Length 291
Best Local Similarity	100.0%	Pred. No. 1.7e-55		
Matches 127	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	GTGGGATACCTGTGCCATCTGCAGGGTCCAGAGTATGGATCCTGCCTTCATGTCACAGC	60	
Db	146	GTGGGATACCTGTGCCATCTGCAGGGTCCAGAGTATGGATCCTGCCTTCATGTCACAGC	205	
QY	61	TGAAAACAGCAGAGGACTGTGTGTGTGGGTGGGGAGAGTAAACCATTCCTTCCACAA	120	
Db	206	TGAAAACAGCAGAGGACTGTGTGTGTGGGTGGGGAGAGTAAACCATTCCTTCCACAA	265	
QY	121	CTGCTGC 127		
Db	266	CTGCTGC 272		

RESULT 3	AA231201	296 bp	RNA	linear	EST 26-FEB-1997
LOCUS	AA231201				
DEFINITION	AA231201.1 Soares mouse 3MNE2.5 Mus musculus cDNA clone IMAGE:673028 5' similar to Wf:ZK287.5 CE0614 ;, mRNA sequence.				
ACCESSION	AA231201				
VERSION	AA231201.1	GI:1853558			
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 296)				
REFERENCE	Warrar,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacey,M., Le,M., Martin,J., Morris,B., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Teising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.				
TITLE	The WashU-HMII Mouse EST Project				
JOURNAL	Unpublished				
COMMENT	Contact: Marra M/Mouse EST Project				

FEATURES
source

```

organism="Mus musculus"
mol_type="mRNA"
strain="C57BL/6J"
db_xref="taxon:10090"
/clone="IMAGE:673028"
sex="unknown"
tissue_type="fetus"
dev_stage="12.5dpb total fetus"
lab_host="DH10B"
clone_lib="Scots mouse 3MNE12.5"

```

BASE COUNT	62 a	74 c	92 g	67 t	1 others
ORIGIN					
Query Match	100.0%;	Score 127;	DB 9;	Length 296;	
Best Local Similarity	100.0%;	Pred. No. 1.7e-55;			
Matches 127;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GTGGCATCTCTGTGTCATCTGCAGGGGTCCAGGTGATGGATCGCCCTTCATGTCAAGC	60		
Db	109	GTGCATACCTCTGTGTCATCTGCAGGGGTCCAGGTGATGGATCGCCCTTCATGTCAAGC	168		
QY	61	TGAACAAGCAAGAGGAGCTGTGTGTGTGTGTGGGAGAGTGAACCATTCCTTCCACAA	120		
Db	169	TGAACAAGCAAGAGGAGCTGTGTGTGTGTGTGGGAGAGTGAACCATTCCTTCCACAA	228		
QY	121	CTGCTGCG	127		
Db	229	CTGCTGCG	235		
RESULT 4					
BE624155					
LOCUS	BE624155	303 bp	mRNA	linear	EST 24-AUG-2000
DEFINITION	u42c01.y1 Soares mammary gland MMLMG Mus musculus cDNA clone				
	IMAGE:3374592.5' similar to WP:R10A10.2 CE12670 ZINC FINGER PROTEIN				
	; mRNA sequence.				
ACCESSION	BE624155				
VERSION	BE624155.1	GI:9904571			
KEYWORDS	EST.				
SOURCE	Mus musculus				
ORGANISM	Mus musculus (house mouse)				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 303)				
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
JOURNAL	Tumor Gene Index				
COMMENT	Unpublished				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapbs-rt@mail.nih.gov				
	This clone is available royalty-free through LNL; contact the				
	IMAGE Consortium (info@image.llnl.gov) for further information.				
	MG1:1084196				
	Seq primer: -40RP from Gibco				
	High quality sequence stop: 292.				
FEATURES	location/Qualifiers				
source	1. 303				
	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/db_xref="taxon:10090"				
	/clone="IMAGE:3374592"				
	/sex="female (lactating)"				
	/tissue_type="mammary gland"				
	/lab_host="DH10B"				
	/clone_idb="Soares mammary gland_MMLMG"				
	/note="Vector: pRT3D-Pac (Pharmacia) with a modified				
	polylinker; 1st strand cDNA was prepared from mammary				
	gland tissue from a lactating female, and was then primed				
	with a Not I - oligo(dT) primer. Double-stranded cDNA was				
	ligated to Eco RI adaptors (Pharmacia), digested with Not				
	I and cloned into the Not I and Eco RI sites of the				
	modified pRT73 vector. Library is normalized. Library				
	was constructed by Bento Soares and M. Fatima Bonaldo."				

[illegible]

Db 193 TGAACAAGCAGAGGACTGTGTGGTGGGAGAGTGTAACCATTCCTCCACAA 252

QY 121 CTGCTGC 127

Db 253 CTGCTGC 259

RESULT 10
LOCUS BE692680
DEFINITION uw44b12.y1 Soares mammary_gland NMLMG Mus musculus cDNA clone
IMAGE:3464831.5, similar to WP:R10A10.2 CE12670 ZINC FINGER PROTEIN
; mRNA sequence.

ACCESSION BE692680
VERSION BE692680
KEYWORDS EST
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 323)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Other ESTs: uw44b12.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1384191
Seq primer: -40RP from Gibco
High quality sequence stop: 260.

FEATURES
source location/Qualifiers
1..323
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3464831"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NMLMG"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pRT3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 73 a 80 c 102 g 68 t

ORIGIN

Query Match 100.0%; Score 127; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 1,7e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCATACCTGTGCATCTGCAGGATCGAGTGATGATGCTTCCTTCATGTCAGC 60
Db 125 GTGCATACCTGTGCATCTGCAGGATCGAGTGATGATGCTTCCTTCATGTCAGC 184

QY 61 TGAACAAGCAGAGGAGAGTGTGTGGTGGGAGAGTGTAACCATTCCTCCACAA 120
Db 185 TGAACAAGCAGAGGAGAGTGTGTGTGGTGGGAGAGTGTAACCATTCCTCCACAA 244

QY 121 CTGCTGC 127

Db 245 CTGCTGC 251

RESULT 11
BF012987

LOCUS BF012987 325 bp mRNA linear EST 29-DEC-2000
DEFINITION ux83b03.y1 McCarrey Eddy type B spermatogonia Mus musculus cDNA
clone IMAGE:3655085.5, similar to TR:Q9WTZ1 Q9WTZ1 ZINC RING FINGER
PROTEIN SAG.; mRNA sequence.

ACCESSION BF012987
VERSION BF012987
KEYWORDS EST
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 325)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person
B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schuck,R., Ritter
B., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1417389
Seq primer: primer name ambiguous
High quality sequence stop: 316.

FEATURES
source location/Qualifiers
1..325
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:3655085"
/sex="male"
/tissue_type="type B spermatogonia, pooled from multiple
mice"
/dev_stage="8 day"
/lab_host="DH10B (phage-resistant)"
/clone_lib="McCarrey Eddy type B spermatogonia"
/note="Organ: testis; Vector: pBluescript SK+ (Stratagene
); Site 1: XhoI; Site 2: EcoRI; cDNA oligo dt-primed
[5'-(GA)10-ACTAGCTCGAGTGTGTTTCTTTT-3'] and directionally
cloned using 5' linkers 5'-AATTCGACAGAG-3' and
5'-CTCGGCGG-3'. Size selection of >400bp material gives
average insert size ranging from 1-2 kb. Library was mass
excised (from lambda-unizap-XR) and resulting
single-stranded phagemids were prepped and transformed
into DH10B. Library contains 96% recombinants.
References: J. Androl. 20:635-639 and Gene 25:263-269.
Library constructed and donated by J. McCarrey, Ph.D.
(Southwest Foundation for Biomedical Research, Dept. of
Genetics); excision done by E.W. Eddy, Ph.D. (National
Institutes of Health, National Institute of Environmental
Health Sciences). Original lambda-based library is
available through ARCC, catalog #63417."

BASE COUNT 72 a 77 c 106 g 70 t

ORIGIN

Query Match 100.0%; Score 127; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 1,7e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCATACCTGTGCATCTGCAGGATCGAGTGATGATGCTTCCTTCATGTCAGC 60
Db 81 GTGCATACCTGTGCATCTGCAGGATCGAGTGATGATGCTTCCTTCATGTCAGC 140

QY 61 TGAACAAGCAGAGGAGAGTGTGTGGTGGGAGAGTGTAACCATTCCTCCACAA 120
Db 141 TGAACAAGCAGAGGAGAGTGTGTGTGGTGGGAGAGTGTAACCATTCCTCCACAA 200

QY 121 CTGCTGC 127
 DB 201 CTGCTGC 207

RESULT 12
 AA623723 328 bp mRNA linear EST 14-OCT-1997
 LOCUS vq72g03.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
 DEFINITION IMAGE:1107892 5' similar to WP:ZK287.5 CE06614 ;, mRNA sequence.
 AA623723
 ACCESSION
 VERSION
 KEYWORDS
 EST. GI:2527559
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 328)
 REFERENCE
 AUTHORS
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished
 JOURNAL
 COMMENT
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LML ; contact the
 IMAGE Consortium (info@image.jnl.gov) for further information.
 MGI:606060
 Possible reversed clone; similarity on wrong strand
 High quality sequence stop: 263.
 Location/Qualifiers
 1. 328
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J x DBA/2J F1"
 /db_xref="taxon:10090"
 /clone="IMAGE:1107892"
 /tissue_type="embryo"
 /dev_stage="2-cell"
 /lab_host="DH10B"
 /clone_lib="Knowles Solter mouse 2 cell"
 /note="Organ: embryo; Vector: pBluescribe (modified);
 Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
 mRNA prepared from 13,500 2-cell stage embryos. Primer:
 SalI (dr): 5'-CGGTCAGCCGTCGACGCTTTTCTTTT-3'. CDNA
 were cloned into the MluI/SalI sites of a modified
 pBluescribe vector using commercial linkers (NEB).
 Average insert size: 1.2 kb."
 BASE COUNT
 ORIGIN
 76 a 78 c 104 g 70 t
 Query Match 100.0%; Score 127; DB 9; Length 328;
 Best Local Similarity 100.0%; Pred. No. 1.7e-55;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 236 CTGCTGC 242

RESULT 13
 BY342015 328 bp mRNA linear EST 12-DEC-2002
 LOCUS BY342015 RIKEN full-length enriched, whole joints Mus musculus cDNA
 DEFINITION clone l230037K21 5', mRNA sequence.
 BY342015
 ACCESSION
 VERSION
 KEYWORDS
 EST. GI:26571503
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 328)
 REFERENCE
 AUTHORS
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Niki, T., Otsu, N., Saito, R., Suzuki, H., Yamashita, T., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gojobori, T., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A.,
 Quackenbush, J., Schmitt, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbett,
 L.E., Cousins, S., Dalla, R., Dragani, T.A., Fletcher, C.F., Forrest,
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
 Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,
 P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perce, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ring,
 Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, D., Ring,
 B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempile, C.A., Setou,
 M., Shimada, K., Sultana, R., Takemura, Y., Taylor, M.S., Teasdale,
 R.D., Tomita, M., Verardo, R., Wagner, L., Whistler, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, T., Alizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
 K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saitama-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Alizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane,
 T., Imotani, K., Iehi, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A.,
 Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
 Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
 M., Waki, K., Watanabe, A., Yamashita, M. and Hayashizaki, Y. Direct
 Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

RESULT 15
 A110259 332 bp mRNA linear EST 02-OCT-1997
 LOCUS mp07g09.r1 Life Tech mouse embryo 8 5dpc 10664019 Mus musculus cDNA
 DEFINITION clone IMAGE:568576 5' similar to WP:ZK287.5 CE06614 ; mRNA
 sequence.
 ACCESSION A110259
 VERSION A110259.1 GI:1662195
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 332)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI: 343224
 Seq primer: -28M13 rev1 from Amersham
 High quality sequence stop: 329.
 Location/Qualifiers
 1..332
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:568576"
 /tissue_type="embryo"
 /dev_stage="8.5dpc embryos"
 /lab_host="DH10B"
 /clone_lib="Life Tech mouse embryo 8 5dpc 10664019"
 /note="Organ: whole embryo; Vector: PCMV-SPORT2; Site: 1:
 SalI; Site 2: NotI; Cloned unidirectionally. Primer:
 Oligo dT: 8.5dpc embryos. PCMV-SPORT2 vector."
 74 a 80 c 108 g 70 t
 BASE COUNT
 ORIGIN
 Query Match 100.0%; Score 127; DB 9; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.7e-55;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GTGCGATCCCTGTCATCTGCGAGGTCGAGTGATGCGCTTCGATGTCAGGC 60
 Db 93 GTGCGATCCCTGTCATCTGCGAGGTCGAGTGATGCGCTTCGATGTCAGGC 152
 Oy 61 TGAAGAAAG 120
 Db 153 TGAAGAAAG 212
 Oy 121 CTGCTGC 127
 Db 213 CTGCTGC 219

RESULT 16
 LOCUS BY337380 346 bp mRNA linear EST 11-DEC-2002
 DEFINITION BY337380 RIKEN full-length enriched, whole joints Mus musculus cDNA
 clone L230006P10 5', mRNA sequence.

ACCESSION BY337380
 VERSION BY337380.1 GI:26532986
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 346)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
 Quackenbush, J., Schiml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisel, K. W., Blake, J. A., Brad, D., Bruscia, V., Chochia, C., Corbani,
 L. E., Cousins, S., Dalla, E., Dreagan, T. A., Fletcher, C. F., Forrest,
 A., Fraser, K. S., Gasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustincich, S., Hitokawa, N., Jackson, I. J.,
 Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M.,
 King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
 P. A., Maglott, D. R., Maltais, L., Matchonni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Petrea, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
 B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C. A., Secou,
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
 R. D., Tomita, M., Verardo, R., Wagner, L., Walstedt, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yangisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sakai, D., Shibata,
 K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-22 Saitama-cho, Tsukuba-shi, Ibaraki, 305-3858, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9226
 Email: genome-res@gs.c.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Aizawa, K., Akimura, T., Arakawa, T.,
 T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A.,
 Murata, M., Nakamura, M., Nomura, K., Numata, K., Ohno, M., Sakai, K.,
 Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
 M., Waki, K., Watanabe, A., Yamashita, M. and Hayashizaki, Y. Direct
 Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multipillar sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in Riken.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Vasilius Aidinis (Biomedical Sciences
 Research Center "Al. Fleming" Institute of Immunology 14-16 Al.
 Fleming street 16672 Vari, Greece) whose assistance we gratefully

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by David A. Hume (Dept. of Biochemistry and Microbiology/Parasitology Institute for Molecular Biocscience University of Queensland Brisbane, Q 4072 Australia) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1. 351
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1830081812"
/tissue_type="bone marrow"
/cell_type="macrophage"
/clone_id="RIKEN full-length enriched, bone marrow macrophage"

BASE COUNT

72 a 93 c 113 g 73 t

Query Match 100.0%; Score 127; DB 13; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGCATCTGTGGCATCTGCAGGCTCAGGTGATGATGCTGCTTCATGATCAAC 60
DB 160 GTGGCATCTGTGGCATCTGCAGGCTCAGGTGATGATGCTGCTTCATGATCAAC 219
QY 61 TAAAAACAAGACAGAGACTGTGTGTGTGTGGGAGAGGTATCAATCTCTTCACAA 120
DB 220 TAAAAACAAGACAGAGACTGTGTGTGTGTGGGAGAGGTATCAATCTCTTCACAA 279
QY 121 CTGGCTGC 127
DB 280 CTGGCTGC 286

RESULT 20 352 bp mRNA linear EST 07-DEC-2002
LOCUS BY091208
DEFINITION BY091208 RIKEN full-length enriched, 13 days embryo whole body Mus musculus cDNA clone K630085B17 5', mRNA sequence.
ACCESSION BY091208

VERSION

BY091208.1 GI:26200367
EST.
Mus musculus (house mouse)

KEYWORDS

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 352)
Oikawa, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikiido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Tagli, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Baladrelli, R., Hall, D.P., Bull, C., Hume, D.A., Quakebush, J., Schmitt, L.M., Kanpin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Bruscic, V., Chochia, C., Cobani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perce, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Borja, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL MEDLINE
PUBMED 22354683
COMMENT 1246851

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source Location/Qualifiers
1. 352

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K630085B17"
/tissue_type="whole body"
/dev_stage="13 days embryo"
/clone_11b="RIKEN full-length enriched, 13 days embryo whole body"

BASE COUNT      80 a      86 c      114 g      72 t

ORIGIN
Query Match      100.0%; Score 127; DB 13; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCGATACCTGTGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTCGATGTCAGC 60
    |||||
DB 139 GTGCGATACCTGTGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTCGATGTCAGC 198
    |||||

QY 61 TGAACAAGCAAGAGAGACTGTGTGTGTGTGTGGGAGAGTGAACCATTCCTCCACAA 120
    |||||
DB 199 TGAACAAGCAAGAGAGACTGTGTGTGTGTGTGGGAGAGTGAACCATTCCTCCACAA 258
    |||||

QY 121 CTGCTGC 127
    |||||
DB 259 CTGCTGC 265
    |||||

RESULT 21
BG089348          355 bp      mRNA      linear      EST 26-JAN-2001
LOCUS             BG089348
DEFINITION        ut6a03.y1 Soares mouse NMGB bcell Mus musculus cDNA clone
IMAGE:333264 5' similar to TR:Q9WTZ1 Q9WTZ1 ZINC RING FINGER
PROTEIN SEQ. ;, mRNA sequence.
BG089348
ACCESSION         BG089348.1 GI:12571911
VERSION           BG089348
KEYWORDS          Mus musculus (house mouse)
SOURCE            EST.
ORGANISM          Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 355)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1076808
Seq primer: -40RP from Gibco
High quality sequence stop: 326.
Location/Qualifiers
1..355
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:333264"
/lab_host="DH10B (phage-resistant)"
/clone_11b="Soares mouse NMGB bcell"
/note="Organ: germinal B-cell; Vector: pRT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCCCTGATTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pRT73 vector.
Library is normalized; constructed by Bento Soares and
M.Patricia Bonaldo."

BASE COUNT      85 a      83 c      111 g      76 t

```

```

ORIGIN
Query Match      100.0%; Score 127; DB 10; Length 355;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCGATACCTGTGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTCGATGTCAGC 60
    |||||
DB 74 GTGCGATACCTGTGCCATCTGCAGGGTCCAGGTGATGATGCTGCTTCGATGTCAGC 133
    |||||

QY 61 TGAACAAGCAAGAGAGACTGTGTGTGTGTGTGGGAGAGTGAACCATTCCTCCACAA 120
    |||||
DB 134 TGAACAAGCAAGAGAGACTGTGTGTGTGTGTGGGAGAGTGAACCATTCCTCCACAA 193
    |||||

QY 121 CTGCTGC 127
    |||||
DB 194 CTGCTGC 200
    |||||

RESULT 22
BY090965          357 bp      mRNA      linear      EST 07-DEC-2002
LOCUS             BY090965
DEFINITION        BY090965 RIKEN full-length enriched, 13 days embryo whole body Mus
musculus cDNA clone K630083012 5', mRNA sequence.
BY090965
ACCESSION         BY090965.1 GI:26200166
VERSION           BY090965
KEYWORDS          Mus musculus (house mouse)
SOURCE            EST.
ORGANISM          Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 357)
Oikazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nakaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaoka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Brack, D., Brusic, V., Chochua, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragan, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kamai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Kongaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavani, W. J., Petrea, G.,
Pesele, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yangisawa,
M., Yang, L., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Nakamura,
A., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, D., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Shiohira-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akiyama, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane

```


T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

```
1. .357
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K630083012"
/tissue_type="whole body"
/dev_stage="13 days embryo"
/clone_lib="RIKEN full-length enriched, 13 days embryo whole body"

```

BASE COUNT

81 a 89 c 115 g 72 t

ORIGIN

Query Match 100.0%; Score 127; DB 13; Length 357;
 Best Local Similarity 100.0%; Pred No. 1.8e-55;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCGATACCTGTGCGATCTGCAGGCTCCAGTGTGATGCTGCTTCGATGTCAGC 60

Db 139 GTGCGATACCTGTGCGATCTGCAGGCTCCAGTGTGATGCTGCTTCGATGTCAGC 198

QY 61 TGAACAAGCAAGAGAGACTGTGTGTGTCTGGGAGAGTGTACCATTCCTCCACAA 120

Db 199 TGAACAAGCAAGAGAGACTGTGTGTGTCTGGGAGAGTGTACCATTCCTCCACAA 258

QY 121 CTGCTGC 127

Db 259 CTGCTGC 265

RESULT 23 BY02366 361 bp mRNA linear EST 11-DEC-2002

LOCUS BY02366 RIKEN full-length enriched, 14.5 days embryo df/df

DEFINITION Rathe's pouches Mus musculus cDNA clone K820003N18 5', mRNA

ACCESSION BY02366

VERSION BY02366

KEYWORDS EST

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 361)

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,

Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,

Gojohori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,

Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Belset, K.W., Blake, J.A., Brad, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gaidarov, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.D., Jarvis, E.D., Karai, A., Kawai, H., Kawasawa, Y., Kedierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lehar, B., Lyons, P.A., Maglott, D.R., Maltas, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pette, G., Pesole, G., Petrovsky, N., Pillai, R., Pontus, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verrado, R., Wagner, L., Waberski, C., Wang, Y., Watanabe, Y., Wells, L., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carrington, P., Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683
 12466851

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@genome.gsc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Michelle Brinkmeier and Sally Camper (Dept. Human Genetics University of Michigan Medical School 4301 MSRB 3 1500 W. Medical Center Dr. Ann Arbor, MI 48109-0638 USA) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

```
1. .361
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="K820003N18"
/tissue_type="Ratke's pouches"
/dev_stage="14.5 days embryo df/df"
/clone_lib="RIKEN full-length enriched, 14.5 days embryo df/df Ratke's pouches"

```


Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)

I-1-722 Sueni-ro-cho, Isurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Fax: 81-45-503-9216

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>

URL: <http://genome.qsc.riken.go.jp/lin11/genome-resdbc/lin11.go.jp/>

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
Location/Qualifiers
1..376

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1830094D24"
/tissue_type="bone marrow"
/cell_type="macrophage"
/clone_id="RIKEN full-length enriched, bone marrow
macrophage"

BASE COUNT 80 a 99 c 122 g 75 t

Query Match 100.0%; Score 127; DB 13; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGCATACCTGTCATCTGCAGGCTCCAGGTGATGATGCTGCTTCGATGTCAGC 60
DB 151 GTGGCATACCTGTCATCTGCAGGCTCCAGGTGATGATGCTGCTTCGATGTCAGC 210
QY 61 TGAACAAGCAG 120
DB 211 TGAACAAGCAG 270
QY 121 CTGCTGC 127
DB 271 CTGCTGC 277

RESULT 30
LOCUS B161786 382 bp mRNA linear EST 10-DEC-2002
DEFINITION B161786 RIKEN full-length enriched, bone marrow macrophage Mus
musculus cDNA clone 1830035A17 5', mRNA sequence.
VERSION B161786
KEYWORDS B161786.1 GI:26298432
SOURCE EST.
ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 382)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chochia, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Flecher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pereira, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramchandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempole, C. A., Secou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Vetraro, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,
M., Yang, L., Yang, L., Yuan, Z., Zavalon, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Maki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation

JOURNAL of 60,770 full-length cDNAs
MEDLINE Nature 420, 563-573 (2002)
PUBMED 22534683
COMMENT 12466851

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
url: <http://genome.gsc.riken.go.jp/>

Alzawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane,
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Maki, K., Wataniki, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission

Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Tissues were provided by David A. Hume (Dep. of Biochemistry
and Microbiology/Parasitology Institute for Molecular Bioscience
University of Queensland Brisbane, Q 4072 Australia) whose
assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES
Location/Qualifiers
1..382

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1830035A17"
/tissue_type="bone marrow"
/cell_type="macrophage"
/clone_id="RIKEN full-length enriched, bone marrow
macrophage"

BASE COUNT 80 a 98 c 126 g 78 t

Query Match 100.0%; Score 127; DB 13; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGCATACCTGTCATCTGCAGGCTCCAGGTGATGATGCTGCTTCGATGTCAGC 60
DB 150 GTGGCATACCTGTCATCTGCAGGCTCCAGGTGATGATGCTGCTTCGATGTCAGC 209
QY 61 TGAACAAGCAG 120
DB 210 TGAACAAGCAG 269
QY 121 CTGCTGC 127
DB 270 CTGCTGC 276

RESULT 31

BY307739 392 bp mRNA linear EST 11-DEC-2002
 LOCUS BY307739 RIKEN full-length enriched, stroma cell Mus musculus cDNA
 clone 1320001102 5', mRNA sequence.
 ACCESSION BY307739
 VERSION BY307739.1 GI:26498076
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 392)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
 Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batilov, S.,
 Beisel, K. W., Blake, J. A., Brad, D., Brusic, V., Chothia, C., Corbani,
 L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
 A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Giesi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
 Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
 King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
 P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G.,
 Pesole, G., Petrovsky, N., Pili, R., Pontius, J. U., Qi, D.,
 Ramchandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
 B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C. A., Setou,
 M., Shimada, K., Sultana, R., Takenaki, Y., Taylor, M. S., Teasdale,
 R. D., Tomita, M., Vitaranda, R., Wagner, L., Wamboldt, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L. G., Wyszewski, A., Yanagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
 K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 JOURNAL NATURE
 MEDLINE 22354683
 PUBMED 12466851
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S., Hirozane,
 T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,
 Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
 Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
 M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct
 Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to

Prepare mouse tissues.
 Tissues were provided by Takashi Ishikawa (Department of Surgery
 2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama
 236-0004 Japan) whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 Location/Qualifiers
 source 1..392
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="1320001102"
 /cell_type="stroma cell"
 /clone_lib="RIKEN full-length enriched, stroma cell"
 80 a 103 c 129 g 80 t
 BASE COUNT
 ORIGIN
 Query Match 100.0%; Score 127; DB 13; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.8e-55;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGGCATACCTGTGCATCTGCAGGATCAGGTGATGATGCTGCTTCATGTCACAC 60
 |||||
 DB 160 GTGGCATACCTGTGCATCTGCAGGATCAGGTGATGATGCTGCTTCATGTCACAC 219
 |||||
 QY 61 TGAACAAACAGCAGGACCTGT 120
 |||||
 DB 220 TGAACAAACAGCAGGACCTGT 279
 |||||
 QY 121 CTGCTGC 127
 |||||
 DB 280 CTGCTGC 286
 |||||
 RESULT 32
 LOCUS CD565301 393 bp mRNA linear EST 11-JUN-2003
 DEFINITION B0492H01-5 N1A Mouse E6.5 whole embryo cDNA library (long) Mus
 musculus cDNA clone N1A:B0492H01 IMAGE:30447540 5', mRNA sequence.
 ACCESSION CD565301
 VERSION CD565301.1 GI:31609194
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 393)
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 Genome Res. 11 (9), 1553-1558 (2001)
 JOURNAL GENOME RESEARCH
 MEDLINE 21429098
 PUBMED 11544199
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@gsun.grc.nia.nih.gov
 Plate: B0492 row: H column: 01
 Seq primer: M13 Reverse
 High quality sequence scop: 393
 POLYA=No.
 Location/Qualifiers
 source 1..393
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="N1A:B0492H01 IMAGE:30447540"
 /tissue_type="E6.5 whole embryo"

Db 155 GTCGATACCTGTGTCATCTGCAGGGTCCAGGTGATGATGTCGCTTCGATGTCACG 214

Qy 61 TGAACAACGACGAGAGACTGTGTGTGTCGGAGAGAGTAAACATTCTTCACAA 120

Db 215 TGAACAACGACGAGAGACTGTGTGTGTCGGAGAGAGTAAACATTCTTCACAA 274

Qy 121 CTGCTGC 127

Db 275 CTGCTGC 261

RESULT 34
BY095585
LOCUS
DEFINITION BY095585 RIKEN full-length enriched, 17 days embryo whole body Mus
musculus cDNA clone K630106W08 5', mRNA sequence.

ACCESSION
VERSION BY095585.1 GI:26206202

KEYWORDS
SOURCE EST.

ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 395)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, U., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schorbach, C.,
Gojobori, T., Balderelli, R., Hill, D. P., Bulc, C., Hume, D. A.,
Queckenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corban,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. D.,
Jarvis, B. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedziarski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, U. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C. A., Setou,
M., Shmada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, J., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wyszynski, B. S., Yamagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavalin, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kisinkawa, T., Kono, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Aizawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE

JOURNAL
MEDLINE
PUBMED 22354683 12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9232
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arahawa, T., Carninci, P., Fukuda, S., Hirozane,
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tegan,
M., Waki, K., Watanishi, A., Watanishi, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
Location/Qualifiers
1. 395
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K630106W08"
/tissue_type="whole body"
/dev_stage="17 days embryo"
/clone_lib="RIKEN full-length enriched, 17 days embryo
whole body"
BASE COUNT 88 a 97 c 128 g 81 t 1 others

Query Match 100.0%; Score 127; DB 13; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGATACCTGTGTCATCTGCAGGGTCCAGGTGATGATGTCGCTTCGATGTCACG 60

Db 138 GTCGATACCTGTGTCATCTGCAGGGTCCAGGTGATGATGTCGCTTCGATGTCACG 197

Qy 61 TGAACAACGACGAGAGACTGTGTGTGTCGGAGAGAGTAAACATTCTTCACAA 120

Db 198 TGAACAACGACGAGAGACTGTGTGTGTCGGAGAGAGTAAACATTCTTCACAA 257

Qy 121 CTGCTGC 127

Db 258 CTGCTGC 264

RESULT 35
BG277656 397 bp mRNA linear EST 21-FEB-2001
BG277656
LOCUS
DEFINITION ux47c07.y1 Soares_NMAX maxillary process Mus musculus cDNA clone
IMAGE:3513229 5' similar to TR:Q9WTZ1 Q9WTZ1 ZINC RING FINGER
PROTEIN SMG. 1, mRNA sequence.

ACCESSION
VERSION BG277656.1 GI:13073177

KEYWORDS
SOURCE EST.

ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 397)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Other ESTs: ux47c07.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1393805
Seq primer: -40RP from Gibco
High quality sequence stop: 366.
Location/Qualifiers


```

source
1. .397
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:351329"
/tissue_type="maxillary process"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="Soares NMMAx maxillary process"
/note="Vector: pTZ19-Pac (Pharmacia) with a modified
polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA
was primed with a Not I, oligo(dT) primer [5'
TGTTCACATCTGAAGTGGAGCGCCGCGCTTTTCTTTTCTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTZ19 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT
94 a 94 c 125 g 84 t
ORIGIN

Query Match 100.0%; Score 127; DB 10; Length 397;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 GTGCGATACCTGTGCCATCTGCAGGGTCCAGGTGATGATGCTTGCCTTCGATGCAAGC 60
DB 82 GGGCATACCTGTGGCATCTGCAGGGTCCAGGTGATGATGCTTGCCTTCGATGCAAGC 141
QY 61 TGAATAACAGCAAGAGGACTGTGTTGTGGTCTGGGGAGAGGTAAACCATTCCTTCACACA 120
DB 142 TGAATAACAGCAAGAGGACTGTGTTGTGGTCTGGGGAGAGGTAAACCATTCCTTCACACA 201
QY 121 CTGCTGC 127
DB 202 CTGCTGC 208

RESULT 36
BP011540 LOCUS 401 bp mRNA linear EST 06-OCT-2000
DEFINITION us40e01.y1 Soares NMMAx brainchial arch Mus musculus cDNA clone
IMAGE:316560 5' similar to TR:Q9WTZ1 Q9WTZ1 ZINC RING FINGER
PROTEIN SAG. ; mRNA sequence.
BP011540
ACCESSION BP011540.1 GI:10711815
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
MUS MUSCULUS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 401)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNI at:
image.llnl.gov/image/html/iresources.shtml

JOURNAL
COMMENT

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
MGI:1065020
Seq primer: -40RP from GIBCO
High quality sequence stop: 394.
1. .401
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

```

	/clone="IMAGE:3169560"	/tissue_type="branchial arches"
	/dev_stage="embryo, 10.5 dpc"	
	/lab_host="DH10B (phage resistant)"	
	/clone_idb="Soares NMBA branchial arch"	
	/note="Vector: pT73d-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'] TGTACCACTGAGTCGAGCGCGGCAGCATTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."	
BASE COUNT	90 a 100 c 130 g 81 t	
ORIGIN		
Query Match	100.0%; Score 127; DB 10; Length 401;	
Best Local Similarity	100.0%; Pred. No. 1.be-55;	
Matches 127; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GGGCATTCCGTGTCATCTGCAGGGTCACAGTGATGATGCCCTTCATGTCAGC 60
Db	134	GTCGCATACCCTGTGCATCTGCAGGGTCACAGTGATGATGCCCTTCATGTCAGC 193
QY	61	TGAACAACAGAAGAGCATCTGTGTGTGGTCTGGGAGAGTGTAACCATTCCTTCACAA 120
Db	194	TGAACAACAGAAGAGCATCTGTGTGTGGTCTGGGAGAGTGTAACCATTCCTTCACAA 253
QY	121	CTGCTGC 127
Db	254	CTGCTGC 260
RESULT 37		
LOCUS	W34101	409 bp mRNA linear EST 11-SEP-1996
DEFINITION	mb01c09.r1 Soares mouse pJNMF19.5 Mus musculus cDNA clone IMAGE:318928 5' similar to WP:F5G12.9 CE00978 ;, mRNA sequence.	
ACCESSION	W34101	
VERSION	W34101.1	GI:1316146
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 409)	
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Thaisring,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.	
TITLE	The WashU-HHMI Mouse EST Project	
JOURNAL	Unpublished	
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watsn.wustl.edu This clone is available royalty-free through LLNL ; contact the IMGE Consortium (info@imge.llnl.gov) for further information. MGI:209544	
FEATURES	Seq primer: ETPrimer High quality sequence stop: 408. Location/Qualifiers 1..409 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /clone="IMAGE:318928" /dev_stage="19.5 dpc total fetus"	

Qy	1	GTGGCAGTACCTGTGCATCTGCAGGGGTCCAGGTATGATGCTGACCTTCATGTCAAGC	60
Db	144	GTCCCATACCTTGTGCATCTGCAGGGTCCAGGTATGATGCTGACCTTCATGTCAAGC	203
Qy	61	TGAATAACAAGCAAGAGGAGCTGTGTGTGTGTGTGTGGAGAGTGTAACCATTCCTTCACAA	120
Db	204	TGAATAACAAGCAAGAGGAGCTGTGTGTGTGTGTGTGGAGAGTGTAACCATTCCTTCACAA	263
Qy	121	CTGCTGC	127
Db	264	CTGCTGC	270
RESULT 41			
LOCUS	BE333776		
DEFINITION	BE333776	427 bp	mRNA
ACCESSION	BE333776		linear
VERSION	BE333776.1	GI:9207552	EST 14-JUL-2000
KEYWORDS	EST.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 427)		
JOURNAL	NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		
	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgsaps-remail.nih.gov		
	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo		
	, Ph.D.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA sequencing by: Washington University Genome Sequencing Center		
	clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	image.llnl.gov/image/html/iresources.shtml		
FEATURES	MGI:1063811		
source	Seg primer: -40RP from Gibco		
	High quality sequence stop: 426.		
	Location/Qualifiers		
	1..427		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:3168351"		
	/tissue_type="branchial arches"		
	/dev_stage="embryo, 10.5 dpc"		
	/lab_host="DH10B (phage resistant)"		
	/clone_1ib="Soares NMEBA branchial arch"		
	/note="Vector: pRTT3D-Pac (Pharmacia) with a modified		
	polylinker. Site 1: NotI; Site 2: EcoRI; 1st strand cDNA		
	polylinker. Site 1: NotI; Site 2: EcoRI; 1st strand cDNA		
	was primed with a NotI oligo(dT) primer [5'		
	TGTTACCAATCTGAAGTGGAGCGCGCCAGCATGATTTTCTTTTCTTTTCTTTTCTTTT		
	3']; double-stranded cDNA was ligated to Eco RI adaptors		
	(Pharmacia), digested with NotI and cloned into the NotI		
	I and Eco RI sites of the modified pRTT3D vector. Library		
	constructed and normalized by Bento Soares and M.Fatima		
	Bonaldo."		
BASE COUNT	100 a	97 c	134 g
ORIGIN			96 t
Query Match	100.0%;	Score 127;	DB 10;
Best Local Similarity	100.0%;	Pred. No. 1.8e-55;	Length 427;
Matches 127;	Conservative 0;	Mismatches 0;	Indels 0;
	Gaps 0;		
Qy	1	GTGGATTCCTGTGCATCTGCAGGGTCCAGGTATGATGCTGACCTTCATGTCAAGC	60

[illegible]

RESULT 43
AA815479
LOCUS
DEFINITION
AA815479 440 bp mRNA linear EST 13-FEB-1998
vp20a07.r1 Soares mammary_gland_NBMWG Mus musculus cDNA clone
IMAGE:1069140 5' similar to TR:P91404 P91404 SIMILARITY TO A3HC4
ZINC FINGER. ; mRNA sequence.

ACCESSION
AA815479
VERSION
AA815479.1 GI:2885075
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 440)

REFERENCE
AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:591500
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 398.

FEATURES
source
1..440
location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1069140"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland_NBMWG"
/note="Organ: mammary gland; Vector: p773D-Pac (Pharmacia
with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer (5'
TGTTCACATCTGAAGTGGAGCGCGCGCATGTGTTTTTTTTTTTTTTTTTTT
T 3'); double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p773 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

BASE COUNT
ORIGIN
90 a 99 c 143 g 108 t

Query Match 100.0%; Score 127; DB 9; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.9e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGATACCTGCTGCATCTGCAGGATCCAGGTATGATGCTTCCTTCATGTCACAGC 60
DB 114 GTGGATACCTGCTGCATCTGCAGGATCCAGGTATGATGCTTCCTTCATGTCACAGC 173
QY 61 TGAACAAGCAG 120
DB 174 TGAACAAGCAG 233
QY 121 CTGCTGC 127
|||||

DB 234 CTGCTGC 240

RESULT 44
AA239898
LOCUS
DEFINITION
AA239898 443 bp mRNA linear EST 03-MAR-1997
mx61h04.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:692791 5'
similar to WP:ZK287.5 CE06614 ; , mRNA sequence.

ACCESSION
AA239898
VERSION
AA239898.1 GI:1863937
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 443)

REFERENCE
AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:426351
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 403.

FEATURES
source
1..443
location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:692791"
/tissue_type="Liver"
/lab_host="DH10B"
/clone_lib="Soares mouse NML"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer (5'
TGTTCACATCTGAAGTGGAGAGCGCGCGCATGTGTTTTTTTTTTTTTTTTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

BASE COUNT
ORIGIN
100 a 106 c 144 g 93 t

Query Match 100.0%; Score 127; DB 9; Length 443;
Best Local Similarity 100.0%; Pred. No. 1.9e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGATACCTGCTGCATCTGCAGGATCCAGGTATGATGCTTCCTTCATGTCACAGC 60
DB 133 GTGGATACCTGCTGCATCTGCAGGATCCAGGTATGATGCTTCCTTCATGTCACAGC 192
QY 61 TGAACAAGCAG 120
DB 193 TGAACAAGCAG 252
QY 121 CTGCTGC 127
DB 253 CTGCTGC 259
|||||

RESULT	45
W34374	
LOCUS	W34374
DEFINITION	mc6d06.r1 Saaree mouse embryo NDMc13.5 14.5 Mus musculus cDNA clone IMAGE:352523 5' similar to WP:F35G12.9 CE00978 ; mRNA sequence.
ACCESSION	W34374
VERSION	W34374.1
KEYWORDS	GI:1316302
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 445) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,U., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R. The WashU-HMNI Mouse EST Project Unpublished Contact: Marra M/Mouse EST project
TITLE	
JOURNAL	
COMMENT	

```

High quality sequence stop: 444.
FEATURES
  Location/Qualifiers
    source             1. .445

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:352523"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/clone_idb="Soares mouse embryo NBME13.5 14.5"
/note="(vector: p17T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I and Eco RI sites of the modified
TGTATCCATCTGAAGTGGAGCGCGCGGGAATTTTTTTTTTTTTTTTTTTT
T 3'] , on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 ] ; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p17T3 vector. Library was through one round of
normalization, and was constructed by Bento Soares and
M.Felina Bonaïdo. "

```

Query Match	100.0%;	Score 127;	DB 14;	Length 445;
Best Local Similarity	100.0%;	Pred. No. 1,9e-55;		
Matches 127;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GTGGGATACCTGTGGCCATCTGCAGGGTCAGAGTGAATGATACCTGCTTCATGTCAAGC	60	
Db	28	GTGGGATACCTGTGGCCATCTGCAGGGTCAGAGTGAATGATACCTGCTTCATGTCAAGC	87	
QY	61	TGAATAACAAGAAAGGACCTGTGTGTGCTCTGGGGAAGTGAACCTTCTCTCCACA	120	
Db	88	TGAATAACAAGAAAGGACCTGTGTGTGCTCTGGGGAAGTGAACCTTCTCTCCACA	147	

Db	148	CTGCTGC	154
121	CTGCTGC	127	
148	CTGCTGC	154	

RESULT	46
LOCUS	AA036356
DEFINITION	447 bp mRNA linear EST 26-AUG-1996
ACCESSION	AA036356
VERSION	AA036356.1
KEYWORDS	GI:1509484
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 447)
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubouque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Teisling,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Maerston,R.
TITLE	The Washu-HHMI Mouse EST Project
JOURNAL	Unpublished
COMMENT	Contact: Marra M/Mouse EST Project

WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through INL, contact the
IMAGS Consortium (info@image.lnl.gov) for further information.
MGI:283101
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 446.

	FEATURES
	source
	location/Qualifiers
	1..447
	/organism="Mus musculus"
	/mol_type="mRNA"
	/db_xref="taxon:10090"
	/clone="IMAGE:472357"
	/dev_stage="19.5 dpc total fetus"
	/lab_host="DH10B (ampicillin resistant)"
	/clone_1lb="Soares mouse p3M19.5"
	/note="Vector: pTV3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dAT) primer [5'] TGTACCAATCTGAAGGAGCGGGCGGCACTTTTTTTTTTTTTT 3'] double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pTV3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. library constructed by Bento Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."
BASE COUNT	101 a
ORIGIN	109 c 144 g 93 t

	Query Match	100.0%;	Score 127;	DB 9;	Length 447;
	Best Local Similarity	100.0%;	Pred. No. 1.9e-55;		
	Matches 127;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	GTGGGATACCTGTGCCATCTGCAAGGGTCCAGGGATGATGATGCTGCTTGGATGTCAAGC	60		
Db	135	GTGGGATACCTGTGCCATCTGCAAGGGTCCAGGGATGATGATGCTGCTTGGATGTCAAGC	194		
Oy	61	TGAAAACAACGACGAGAGACTGTGTGTGGCTGGGGAGAGTGAACCATTCCTCCACAA	120		
Db	195	TGAAAACAACGACGAGAGACTGTGTGTGGCTGGGGAGAGTGAACCATTCCTCCACAA	254		
Oy	121	CTGCTGC	127		

Db 255 CTGCTGC 261

|||||

RESULT 47
BE854114
LOCUS
DEFINITION

447 bp mRNA linear EST 27-SEP-2000
ux27009.y1 Soares_NMAX maxillary process mus musculus cDNA clone
IMAGE:3511552.5' similar to TR:Q9WTZ1 Q9WTZ1 ZINC RING FINGER
PROTEIN SAG. ; mRNA sequence.

ACCESSION
BE854114
VERSION
BE854114.1 GI:103312726
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 447)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

JOURNAL
Tumor Gene Index

COMMENT
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1394432
Seq primer: -40RP from Gibco
High quality sequence stop: 434.
Location/Qualifiers
1. .447
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3511552"
/tissue_type="maxillary process"
/lab_host="DH10B (phage-resistant)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: NotI; Site_2: BcoRI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5']
TGTTACCATCTGAAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT
108 a 100 c 138 g 101 t

ORIGIN

Query Match 100.0%; Score 127; DB 10; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.9e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGATACCTGTGCATCTGCAGGGTCCAGGTGATGATGCTGCTTCATGATCAAGC 60
|||||
DB 109 GTGCGATACCTGTGCATCTGCAGGGTCCAGGTGATGATGCTGCTTCATGATCAAGC 168
|||||

QY 61 TGAACAACAAGCAAGAGACTGTGTGTGTGCTGGGAGAGATGAACATTCCTTCCACAA 120
|||||
DB 169 TGAACAACAAGCAAGAGACTGTGTGTGTGCTGGGAGAGATGAACATTCCTTCCACAA 228
|||||

QY 121 CTGCTGC 127
|||||
DB 229 CTGCTGC 235

RESULT 48
BF152599
LOCUS
DEFINITION

450 bp mRNA linear EST 29-DEC-2000
uz35e03.y1 NCI CGAP Mam5 mus musculus cDNA clone IMAGE:3671068.5'
similar to TR:Q9WTZ1 Q9WTZ1 ZINC RING FINGER PROTEIN SAG. ; mRNA
sequence.

ACCESSION
BF152599
VERSION
BF152599.1 GI:11033994
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 450)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

JOURNAL
Tumor Gene Index

COMMENT
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
image.llnl.gov/image/html/iresources.shtml
MGI:1431836
Seq primer: -40RP from Gibco
High quality sequence stop: 436.
Location/Qualifiers
1. .450
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3671068"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT
104 a 102 c 136 g 107 t 1 others

ORIGIN

Query Match 100.0%; Score 127; DB 10; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.9e-55;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGATACCTGTGCATCTGCAGGGTCCAGGTGATGATGCTGCTTCGATGTCAAGC 60
|||||
DB 89 GTGCGATACCTGTGCATCTGCAGGGTCCAGGTGATGATGCTGCTTCGATGTCAAGC 148
|||||

QY 61 TGAACAACAAGAGAGACTGTGTGTGTGCTGGGAGAGATGAACATTCCTTCCACAA 120
|||||
DB 149 TGAACAACAAGAGAGACTGTGTGTGTGCTGGGAGAGATGAACATTCCTTCCACAA 208
|||||

QY 121 CTGCTGC 127
|||||
DB 209 CTGCTGC 215

RESULT 49
BY242401
LOCUS
DEFINITION

452 bp mRNA linear EST 10-DEC-2002
BY242401 RIKEN full-length enriched, visual cortex Mus musculus
cDNA clone K230034E14.5', mRNA sequence.

ACCESSION
BY242401
VERSION
BY242401.1 GI:26423913
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT	
1 (bases 1 to 452)	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bulic, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Petlea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyszaw-Boris, A., Yanagisawa, M., Yang, T., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Nature 420, 563-573 (2002)				
		1246681					
		1246683					
		1246685					
		1246687					
		1246689					
		1246691					
		1246693					
		1246695					
		1246697					
		1246699					
		1246701					
		1246703					
		1246705					
		1246707					
		1246709					
		1246711					
		1246713					
		1246715					
		1246717					
		1246719					
		1246721					
		1246723					
		1246725					
		1246727					
		1246729					
		1246731					
		1246733					
		1246735					
		1246737					
		1246739					
		1246741					
		1246743					
		1246745					
		1246747					
		1246749					
		1246751					
		1246753					
		1246755					
		1246757					
		1246759					
		1246761					
		1246763					
		1246765					
		1246767					
		1246769					
		1246771					
		1246773					
		1246775					
		1246777					
		1246779					
		1246781					
		1246783					
		1246785					
		1246787					
		1246789					
		1246791					
		1246793					
		1246795					
		1246797					
		1246799					
		1246801					
		1246803					
		1246805					
		1246807					
		1246809					
		1246811					
		1246813					
		1246815					
		1246817					
		1246819					
		1246821					
		1246823					
		1246825					
		1246827					
		1246829					
		1246831					
		1246833					
		1246835					
		1246837					
		1246839					
		1246841					
		1246843					
		1246845					
		1246847					
		1246849					
		1246851					
		1246853					
		1246855					
		1246857					
		1246859					
		1246861					
		1246863					
		1246865					
		1246867					
		1246869					
		1246871					
		1246873					
		1246875					
		1246877					
		1246879					
		1246881					
		1246883					
		1246885					
		1246887					
		1246889					
		1246891					
		1246893					
		1246895					
		1246897					
		1246899					
		1246901					
		1246903					
		1246905					
		1246907					
		1246909					
		1246911					
		1246913					
		1246915					
		1246917					
		1246919					
		1246921					
		1246923					
		1246925					
		1246927					
		1246929					
		1246931					
		1246933					
		1246935					
		1246937					
		1246939					
		1246941					
		1246943					
		1246945					
		1246947					
		1246949					
		1246951					
		1246953					
		1246955					
		1246957					
		1246959					
		1246961					
		1246963					
		1246965					
		1246967					
		1246969					
		1246971					
		1246973					
		1246975					
		1246977					
		1246979					
		1246981					
		1246983					
		1246985					
		1246987					
		1246989					
		1246991					
		1246993					
		1246995					
		1246997					
		1246999					
		1247001					
		1247003					
		1247005					
		1247007					
		1247009					
		1247011					
		1247013					
		1247015					
		1247017					
		1247019					
		1247021					
		1247023					
		1247025					
		1247027					
		1247029					
		1247031					
		1247033					
		1247035					
		1247037					
		1247039					
		1247041					
		1247043					
		1247045					
		1247047					
		1247049					
		1247051					
		1247053					
		1247055					
		1247057					
		1247059					
		1247061					
		1247063					
		1247065					
		1247067					
		1247069					
		1247071					
		1247073					
		1247075					
		1247077					
		1247079					
		1247081					
		1247083					
		1247085					
		1247087					
		1247089					
		1247091					
		1247093					
		1247095					
		1247097					
		1247099					
		1247101					
		1247103					
		1247105					
		1247107					
		1247109					
		1247111					
		1247113					
		1247115					
		1247117					
		1247119					
		1247121					
		1247123					
		1247125					
		1247127					
		1247129					
		1247131					
		1247133					
		1247135					
		1247137					
		1247139					
		1247141					
		1247143					
		1247145					
		1247147					
		1247149					
		1247151					
		1247153					
		1247155					
		1247157					
		1247159					
		1247161					
		1247163					
		1247165					
		1247167					
		1247169					
		1247171					
		1247173					
		1247175					
		1247177					
		1247179					
		1247181					
		1247183					
		1247185					
		1247187					
		1247189					
		1247191					
		1247193					
		1247195					
		1247197					
		1247199					
		1247201					
		1247203					
		1247205					
		1247207					
		1247209					
		1247211					
		1247213					
		1247215					
		1247217					
		1247219					

Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane,
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Tissues were provided by Michela Fagioli and Takao K. Hensch (
Laboratory for Neuronal Circuit Development Brain Science Institute
RIKEN 2-1 Hirosawa, Wako-shi, Saitama 351-0198 Japan) whose
assistance we gratefully acknowledge. Please visit our web site
(http://genome.gsc.riken.go.jp) for further details.

FEATURES

source

```
1..458
/location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K330010J10"
/tissue_type="visual cortex"
/clone_idb="RIKEN full-length enriched, visual cortex"
BASE COUNT      101 a      114 c      149 g      94 t
ORIGIN
```

Query Match 100.0%; Score 127; DB 13; Length 458;

Best Local Similarity 100.0%; Pred. NO. 1.9e-55;

Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 GTGCAATACCTGTGTCATCTGCAGGTCAGGTGATGATGCTGCTTCGATGTCACAGC 60
      |||
DB      150 GTGCAATACCTGTGTCATCTGCAGGTCAGGTGATGATGCTGCTTCGATGTCACAGC 209
      |||

QY      61 TGAATAACAAGACAAGAGACTGTGTGTGTGCTGTGGGAGAGGTAAACATTCCTTCACAA 120
      |||
DB      210 TGAATAACAAGACAAGAGACTGTGTGTGTGCTGTGGGAGAGGTAAACATTCCTTCACAA 269
      |||

QY      121 CTGCTGC 127
      |||
DB      270 CTGCTGC 276
```

Search completed: November 7, 2003, 11:53:53
Job time : 1303.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 11:59:37 / Search time 1900.92 Seconds
(without alignments)
9640.351 Million cell updates/sec

Title: US-09-509-779-3

Perfect score: 754

Sequence: 1 ATGGCCGACGTGAGACGCG.....CAATTAAGTGCAGTTTAA 754

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estc1:*
10: gb_estc2:*
11: gb_hic:*
12: gb_estc3:*
13: gb_estc4:*
14: gb_estc5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pin:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rdg:*
26: em_gse_pbg:*
27: em_gse_vrt1:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	689	91.4	1201	9	AL547435
2	683.6	90.7	853	13	BQ876591 AGENCOURT
3	681	90.3	1201	13	BX324742 BX324742
4	664.8	88.2	949	12	BG769106 BG769106

5	661.6	87.7	797	12	B1832643	B1832643 603082021
6	659.8	87.5	918	12	B1759082	B1759082 603042879
7	649	86.1	822	12	B1668735	B1668735 603293174
8	647.4	85.9	818	12	B1457840	B1457840 603198212
9	644.4	85.5	805	10	BG708518	BG708518 602670411
10	643	85.3	856	12	B1601855	B1601855 60244729
11	640.2	84.9	754	12	BG766992	BG766992 602740511
12	635.4	84.3	1016	12	BMS56683	BMS56683 602287310
13	634.4	84.1	706	12	BH704660	BH704660 UI-E-C11-
14	634.4	84.1	724	12	B1226556	B1226556 602951527
15	632.6	83.9	922	13	BUI92094	BUI92094 AGENCOURT
16	624.6	82.8	855	10	BE747000	BE747000 601580743
17	623.6	82.7	893	10	BG037017	BG037017 602287310
18	619.4	82.1	1201	13	BX387707	BX387707 EX387707
19	616.2	81.7	797	12	B1858307	B1858307 603384080
20	612.8	81.3	983	13	BQ648608	BQ648608 AGENCOURT
21	609	80.8	856	10	BG753323	BG753323 602731740
22	608.8	80.7	767	12	B1601470	B1601470 603249053
23	607.4	80.6	702	9	A1338342	A1338342 qg96603.x
24	606.6	80.5	847	12	BG773490	BG773490 602720212
25	606	80.4	748	10	BG576768	BG576768 602599091
26	603.4	80.0	782	14	CB990062	CB990062 AGENCOURT
27	597.2	79.2	904	10	BG742338	BG742338 602631707
28	595.8	79.0	692	14	CA449255	CA449255 UI-E-ED0-
29	594.2	78.8	716	12	BG761806	BG761806 602718011
30	592.8	78.6	840	10	BG037022	BG037022 602287341
31	589.6	78.2	827	12	B1828930	B1828930 603075092
32	588.8	78.1	675	12	B1858784	B1858784 603386387
33	587.2	77.9	947	10	BG111792	BG111792 602285379
34	582.4	77.2	629	13	BQ632594	BQ632594 1125C04.X
35	582.4	77.2	745	10	BG121625	BG121625 602351581
36	581.8	77.2	734	10	BE397308	BE397308 601288640
37	573.2	76.3	864	12	B1757881	B1757881 603030486
38	574.8	76.2	793	10	BG716023	BG716023 602677016
39	571.8	75.8	651	10	BG714665	BG714665 602677016
40	566.4	75.1	641	13	BQ632301	BQ632301 1125C04.X
41	559.4	74.2	678	13	BUI601136	BUI601136 AGENCOURT
42	559.2	74.2	946	14	CD556579	CD556579 AGENCOURT
43	558.8	74.1	622	14	CB217926	CB217926 NISC md05
44	558.2	74.0	656	13	BUT30720	BUT30720 UI-E-C11-
45	555.4	73.7	595	10	BF033587	BF033587 601453564

ALIGNMENTS

RESULT 1
AL547435
LOCUS
DEFINITION
AL547435 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CSOD1011YM21 5-PRIME, mRNA sequence.
ACCESSION
AL547435
VERSION
AL547435.2 GI:31269266
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12881506.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqrefgenoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2013.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOD1011YG110P1cluster=2013.f. Contact :
Feng Liang Email : fliang@infotech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS01011AG110P1.

FEATURES

Location/Qualifiers

source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01011M21"
/tissue="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT

461 a 180 c 310 g 211 t 39 others

ORIGIN

Query Match 91.4%; Score 689; DB 9; Length 1201;

Best Local Similarity 97.5%; Pred. No. 4e-183;

Matches 742; Conservative 1; Mismatches 11; Indels 7; Gaps 4;

1 ATGCCGACGTGAAAGACGAGAGAGAAACCTGCGCCCTCTCACTCCGGAGCTCA 60
81 ATGCCGACGTGAAAGACGAGAGAGAAACCTGCGCCCTCTCACTCCGGAGCTCA 140
61 GGCTCCAAAGTCGGGAGGCGACAAAGATGTTCTCCCAAGAAATGGAACGGGTGCCATG 120
141 GGCTCCAAAGTCGGGAGGCGACAAAGATGTTCTCCCAAGAAATGGAACGGGTGCCATG 200
121 TGGAGCTGGGACGTGAGAGTCGATACGTCGCCATCTGACGGGTCCAGGTGATGATGCC 180
201 TGGAGCTGGGACGTGAGAGTCGATACGTCGCCATCTGACGGGTCCAGGTGATGATGCC 260
181 TGTCTTAAGTGTCAAGCTGAAACAAACAGAGACCTGTTGTGTCGCGAGAAATGT 240
261 TGTCTTAAGTGTCAAGCTGAAACAAACAGAGACCTGTTGTGTCGCGAGAAATGT 320
241 AATCATTCCTTCCACAACCTGTCGATGTCCTGTCGGTGGGAAACAGAACATGCGCCCT 300
321 AATCATTCCTTCCACAACCTGTCGATGTCCTGTCGGTGGGAAACAGAACATGCGCCCT 380
301 CTCTCCAGCAGACGCTGGTGTGTCACAAAGATCGGCAATGAGAGTGTGTAAGAGCTTC 360
381 CTCTCCAGCAGACGCTGGTGTGTCACAAAGATCGGCAATGAGAGTGTGTAAGAGCTTC 440
361 TTAAGCGAGTGTTCAGAGCCCTGTCGATCTTGAATCCAGTGCCTTACAAAGGCTTGA 420
441 TTAAGCGAGTGTTCAGAGCCCTGTCGATCTTGAATCCAGTGCCTTACAAAGGCTTGA 500
421 ACATACAGGGGATGAATTTCTCAATAGAGAGCGATGATGTCGTC--TTTGGAGCTC 478
501 ACATACAGGGGATGAATTTCTCAATAGAGAGCGATGATGTCGTCGTCGAGCTC 560
479 ATCAAAGCCTTGTT--AGCATTTGTCAATTTTATCTTCAGAAATTTCTGTGATTAAGA 536
561 ATCAAAGCCTTGTTAGCATTTTGTCAATTTTATCTTCAGAAATTTCTGTGATTAAGA 620
537 AGATAATTATTAAGAGTGTCTTCTCACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 596
621 AGATAATTATTAAGAGTGTCTTCTCACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 680
597 GAAGTGTATTAAG 655
681 GAAGTGTATTAAG 740
656 TATAACAAGGACGTGAGAGAGTTTC--GAGACTTTTTCGATGCTTATGTTGATGATCAGT 713
741 TATAACAAGGACGTGAGAGAGTTTCGAGAGAGTTTTCGATGCTTATGTTGATGATCAGT 800
714 TATAACAAGAGATGTTACATTAACAATTAAGTGCAGTTTAAA 754
801 TATAACAAGAGATGTTACATTAACAATTAAGTGCAGTTTAAA 841

RESULT 2

BO876591

LOCUS

DEFINITION

AGNCOURT 8584244 lupski_sympathetic_trunk Homo sapiens cDNA clone

IMAGE:6192819 5', mRNA sequence.

BO876591

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BO876591 853 bp mRNA linear EST 16-AUG-2002
AGNCOURT 8584244 lupski_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6192819 5', mRNA sequence.
BO876591.1 GI:22268599
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 853)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM13595 row: h column: 04
High quality sequence stop: 608.
Location/Qualifiers
1..853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6192819"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult 16 yr"
/lab_host="DH10B"
/clone_lib="lupski_sympathetic trunk"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGGCGTCCG-3' and
5'-GACTAGTTCATGATGCGAGCGGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

FEATURES

source

BASE COUNT 221 a 191 c 225 g 205 t 1 others
ORIGIN
Query Match 90.7%; Score 683.6; DB 13; Length 853;
Best Local Similarity 97.1%; Pred. No. 1.1e-181;
Matches 739; Conservative 0; Mismatches 15; Indels 7; Gaps 4;
1 ATGCCGACGTGAAAGACGAGAGAGAAACCTGCGCCCTCTCACTCCGGAGCTCA 60
51 ATGCCGACGTGAAAGACGAGAGAGAAACCTGCGCCCTCTCACTCCGGAGCTCA 110
61 GGCTCCAAAGTCGGGAGGCGACAAAGATGTTCTCCCAAGAAATGGAACGGGTGCCATG 120
111 GGCTCCAAAGTCGGGAGGCGACAAAGATGTTCTCCCAAGAAATGGAACGGGTGCCATG 170
121 TGGAGCTGGGACGTGAGAGTCGATACGTCGCCATCTGACGGGTCCAGGTGATGATGCC 180
171 TGGAGCTGGGACGTGAGAGTCGATACGTCGCCATCTGACGGGTCCAGGTGATGATGCC 230
181 TGTCTTAAGTGTCAAGCTGAAACAAACAGAGACCTGTTGTGTCGCGAGAAATGT 240
231 TGTCTTAAGTGTCAAGCTGAAACAAACAGAGACCTGTTGTGTCGCGAGAAATGT 290
241 AATCATTCCTTCCACAACCTGTCGATGTCCTGTCGGTGGGAAACAGAACATGCGTCCT 300

```

Db      291 AATCATCTCTCCACCAACTCTGATGTCCTCTGGGTGAAACAGAACATCGTGCCCT 350
Qy      301 CTGCGCAGAGAGACCTGGGTGTCCTCAAGAAATGCGGAAATGAGAGTGTGAAGGCTTC 360
Db      351 CTGCGCAGAGAGACCTGGGTGTCCTCAAGAAATGCGGAAATGAGAGTGTGAAGGCTTC 410
Qy      361 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTGTATCCAGTGCCCTCAAGAGCTTGA 420
Db      411 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTGTATCCAGTGCCCTCAAGAGCTTGA 470
Qy      421 ACACTACAGGGGATGAATTTCTCAATAGAGCCGATGATCTGTGTC--TTTGAATC 478
Db      471 ACACTACAGGGGATGAATTTCTCAATAGAGCCGATGATCTGTGTC--TTTGAATC 530
Qy      479 ATCAAGCCCTGTGTT--AGCATTTGTCAGTTTATCTTCAGAAATTTCTGTGATTAAGA 536
Db      531 ATCAAGCCCTGTGTTAGCATTTTGTGATCTTCAGAAATTTCTGTGATTAAGA 590
Qy      537 AGATAATTTTAAAGGTGTCCTTCTTACCTGTGTGTGTGTCGACGACAGCTTA 596
Db      591 AGATAATTTTAAAGGTGTCCTTCTTACCTGTGTGTGTGTCGACGACAGCTTA 650
Qy      597 GAAGTGTATTAAGAAAGAGAGCTCCAAATTTGAATCAC--TTATATTTTACCATTTC 655
Db      651 GAAGTGTATTAAGAAAGAGAGCTCCAAATTTGAATCAC--TTATATTTTACCATTTC 710
Qy      656 TATACAAACGAGCTGGAAGCACTTC--GAGACTTTTGTGATGCTTATGTGATCAGT 713
Db      711 TATACAAACGAGCTGGAAGCACTTC--GAGACTTTTGTGATGCTTATGTGATCAGT 770
Qy      714 TAAAAAAGATGTTACAGTAACTAAATTAAGTGCAGTTTAA 754
Db      771 TAAAAAAGATGTTACAGTAACTAAATTAAGTGCAGTTTAA 811

```

```

RESULT 3
LOCUS    BX324742      1201 bp      mRNA      linear      EST 01-MAY-2003
DEFINITION BX324742 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0D1037YP13 5-PRIME, mRNA sequence.
ACCESSION BX324742
VERSION   BX324742.1 GI:30307404
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Li, W.B., Gruber, C., Jesse, J. and Polyes, D.
           Full-length cDNA libraries and normalization
           unpublished
COMMENT   Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: sequef@genoscope.cns.fr Web : www.genoscope.cns.fr
           Library was constructed by Life Technologies, a division of
           Invitrogen. This sequence belongs to sequence cluster 2013.f For
           more information about this cluster, see
           http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0A1037CH07QPI
           &cluster=2013.f. Contact : Feng Liang Email : fliang@lifetech.com
           URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS0A1037CH07QPI.
FEATURES
         source
         1..1201
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="CS0D1037YP13"
         /issue_type="PLACENTA COT 25-NORMALIZED"
         /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
         /note="1st strand cDNA was primed with a NotI-oligo(dT)
         primer. Five prime end enriched, double-strand cDNA was

```

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized.

Query Match 90.3%; Score 681; DB 13; Length 1201;
Best Local Similarity 96.5%; Pred. No. 7, 2e-181;
Matches 734; Conservative 4; Mismatches 16; Indels 7; Gaps 4;

```

BASF COUNT 347 a 182 c 300 g 264 t 108 others
ORIGIN
Qy      1 ATGGCCGAGGTGGAACGAGAGAAACCTGCGCCCTGCTCTACTCCGGAGCTCA 60
Db      81 ATGGCCGAGGTGGAACGAGAGAAACCTGCGCCCTGCTCTACTCCGGAGCTCA 140
Qy      61 GGCTCAACCTGGGAGGCGCAAGAGTCTCTCCCTCAAGAGTGAACCGCGGCGCATG 120
Db      141 GGCTCAACCTGGGAGGCGCAAGAGTCTCTCCCTCAAGAGTGAACCGCGGCGCATG 200
Qy      121 TGGAGCTGGGACGTGAGTGCATACGTGCGCATCTGACAGGGTCCAGGTGATGATGCC 180
Db      201 TGGAGCTGGGACGTGAGTGCATACGTGCGCATCTGACAGGGTCCAGGTGATGATGCC 260
Qy      181 TGTCTTAGATGTCAGCTGAACCAAAACAAGAGACTGTGTTGTGCTGGGAGATGT 240
Db      261 TGTCTTAGATGTCAGCTGAACCAAAACAAGAGACTGTGTTGTGCTGGGAGATGT 320
Qy      241 AATCATCTCTTCCACAACTGTCGATGTCCTCTGTGGTGAACACAAACATGCTCCCT 300
Db      321 AATCATCTCTTCCACAACTGTCGATGTCCTCTGTGGTGAACACAAACATGCTCCCT 380
Qy      301 CTCTGCCACAGGACTGGGTGTGCCAAAGATCGCAATGAGAGTGTGAAGAGCTTC 360
Db      381 CTCTGCCACAGGACTGGGTGTGCCAAAGATCGCAATGAGAGTGTGAAGAGCTTC 440
Qy      361 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTTGATATCAGTCCCTTCAAGAGCTTGA 420
Db      441 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTTGATATCAGTCCCTTCAAGAGCTTGA 500
Qy      421 ACACTACAGGGGATGAATTTCTCAATAGAGCCGATGATCTGTGTC--TTTGAATC 478
Db      501 ACACTACAGGGGATGAATTTCTCAATAGAGCCGATGATCTGTGTC--TTTGAATC 560
Qy      479 ATCAAGCCCTGTGTT--AGCATTTGTCAGTTTATCTTCAGAAATTTCTGTGATTAAGA 536
Db      561 ATCAAGCCCTGTGTTAGCATTTTGTGCACTTTATCTTCAGAAATTTCTGTGATTAAGA 620
Qy      537 AGATAATTTTAAAGGTGTCCTTCTTACCTGTGTGTGTGTGTCGACAGCACTTA 596
Db      621 AGATAATTTTAAAGGTGTCCTTCTTACCTGTGTGTGTGTGTCGACAGCACTTA 680
Qy      597 GAAGTGTATTAAGAAAGAGAGCTCCAAATTTGAATCAC--TTATATTTTACCATTTC 655
Db      681 GAAGTGTATTAAGAAAGAGAGCTCCAAATTTGAATCAC--TTATATTTTACCATTTC 740
Qy      656 TATACAAACGAGCTGGAAGCACTTC--GAGACTTTTGTGATGCTTATGTGATCAGT 713
Db      741 TATACAAACGAGCTGGAAGCACTTC--GAGACTTTTGTGATGCTTATGTGATCAGT 800
Qy      714 TAAAAAAGATGTTACAGTAACTAAATTAAGTGCAGTTTAA 754
Db      801 TAAAAAAGATGTTACAGTAACTAAATTAAGTGCAGTTTAA 841

```

```

RESULT 4
LOCUS    BG769106      949 bp      mRNA      linear      EST 15-MAY-2001
DEFINITION BG769106 602743362F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4873095 5',
            mRNA sequence.
ACCESSION BG769106
VERSION   BG769106.1 GI:14079759
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
1 (bases 1 to 949)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS
NIH-MGC <http://mgi.nci.nih.gov/>
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCPD/DTF

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L10CM1750 row: k column: 16
High quality sequence stop: 748.
Location/Qualifiers

FEATURES
source

1..949
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4873095"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 49"
/note="Organ: skin; Vector: pOT7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT
272 a 260 g 213 t

Query Match 88.2%; Score 664.8; DB 12; Length 949;
Best Local Similarity 97.3%; Pred. No. 2.4e-176;
Matches 730; Conservative 0; Mismatches 12; Indels 8; Gaps 5;

13 GAAGACGAGAGGAAACCTGCGCCCTGCGCTCTCACTCCGGAGCTCAGGCTCCAAGTCG 72
1 GAAGACGAGAGGAAACCTGCGCCCTGCGCTCTCACTCCGGAGCTCAGGCTCCAAGTCG 60
73 GGAGCGGACAGATGTTCTCCTCAAGAGTGAACGCGGTGGCCATGTGAGCTGGAGC 132
61 GGAGCGGACAGATGTTCTCCTCAAGAGTGAACGCGGTGGCCATGTGAGCTGGAGC 120
133 GTGAGTGGATACGTGCGCCATCTGCGAGGGTCCAGGTGATGAGTCCCTTATGATGT 192
121 GTGAGTGGATACGTGCGCCATCTGCGAGGGTCCAGGTGATGAGTCCCTTATGATGT 180
193 CAAGCTGAAACCAACAGAGAGAGCTGTGTGTGCTGTGGGAGATGTATCATCTTC 252
181 CAAGCTGAAACCAACAGAGAGAGCTGTGTGTGCTGTGGGAGATGTATCATCTTC 240
253 CACAACCTGCTGATGCTCCTGTGGGTGAACAGAACATCGTGGCCCTCTGCCAGAG 312
241 CACAACCTGCTGATGCTCCTGTGGGTGAACAGAACATCGTGGCCCTCTGCCAGAG 300
313 GACTGGGTGTCCAAAGAGTGGCAATGAGAGTGTAGAAAGCTTCTTACGGCAGTTG 372
301 GACTGGGTGTCCAAAGAGTGGCAATGAGAGTGTAGAAAGCTTCTTACGGCAGTTG 360
373 TTCAGAGCCCTGTGATCTTGTATCCAGTGCCTTCAAGAGGCTAGAACACTACAGGG 432
361 TTCAGAGCCCTGTGATCTTGTATCCAGTGCCTTCAAGAGGCTAGAACACTACAGGG 420
433 ATGATTTCTTAAATAGAGCCGATGATCTGTGTGTC--TTTGAATATCAAGGCTTG 490
421 ATGATTTCTTAAATAGAGCCGATGATCTGTGTGTCCTTTGGGACTCATCAAGGCTTG 480

QY 491 GTT--AGCATTTGTCAGTTTATCTTCGAAATTTCTGTGATTAAGAGATTAATTTAT 548
DB 481 GTTACAGATTTGTAGTTTATCTTCGAAATTTCTGCATTAAGAGATTAATTTAT 540
QY 549 AAAGTGTCTCTTCTTCACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 608
DB 541 AAAGTGTCTCTTCTTCACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 609 AA-AGGAAAGAGCTCCAAATTAATCACC-TTATTAATTAACCAATTTCTATACACAG 666
DB 601 AACAGGAAAGAGCTCCAAATTAATCACC-TTATTAATTAACCAATTTCTATACACAG 660
QY 667 CAGTGAAGCAGTTTC--GAGCTTTTTCATGCTTATGTTGATGTTGATTAAGAGAT 724
DB 661 CAGTGAAGCAGTTTCAGAGAACTTTTGCAATGCTTATGTTGATGTTGATTAAGAGAT 720
QY 725 GTTACAGTAAACAAATTAAGTCACTTTAA 754
DB 721 GTTACAGTAAACAAATTAAGTCACTTTAA 750

RESULT 5
BI832643
LOCUS
DEFINITION
603082021F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5221325 5',
mRNA sequence.
ACCESSION
BI832643
VERSION
BI832643.1 GI:15944193
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 797).
AUTHORS
NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L10M1556 row: 1 column: 06
High quality sequence stop: 795.
Location/Qualifiers

FEATURES
source

1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5221325"
/lab_host="DH10B"
/clone_lib="NIH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector:
pCW-SpOxT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dt primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."

BASE COUNT
208 a 179 c 212 g 198 t

ORIGIN

Query Match 87.7%; Score 661.6; DB 12; Length 797;
Best Local Similarity 96.9%; Pred. No. 1.8e-175;
Matches 740; Conservative 0; Mismatches 14; Indels 10; Gaps 6;
1 ATGGCCGACGTGAAGACGAGAGGAAACCTGCGCCCTGCGCTCTCACTCCGGAGCTCA 60

```

Db      31 ATGGCCGAGCTGGAGAACGAGAGAAACCTGGCCCTCTCAGCTCCGGAGCTCA 90
Qy      61 GGCTCAAGTCGGGAGAGCCAGAGATGTTCTCCCTCAAGAGTGAACCGGTGGCAG 120
Db      91 GGCTCAAGTCGGGAGAGCCAGAGATGTTCTCCCTCAAGAGTGAACCGGTGGCAG 150
Qy      121 TGGAGCTGGGACGTGAGTGCAGTACGTGCGCATCTGCAAGGATCCAGGTGATGATGCC 180
Db      151 TGGAGCTGGGACGTGAGTGCAGTACGTGCGCATCTGCAAGGATCCAGGTGATGATGCC 210
Qy      181 TGTCTTAGATGTCAGCTGAAAAAACAAGAGAGCTGTGTTGTGCTGGGAGAGATG 240
Db      211 TGTCTTAGATGTCAGCTGAAAAAACAAGAGAGCTGTGTTGTGCTGGGAGAGATG 270
Qy      241 AATCATCTCTCCACAGCTGCTGCAATGCTCTGTGGTGAACAGAACTGCTGCTC 300
Db      271 AATCATCTCTCCACAGCTGCTGCAATGCTCTGTGGTGAACAGAACTGCTGCTC 330
Qy      301 CTCTGCGACAGAGCTGGGTGCTCAAGAGATCGGCAATGAGAGTGTAGAGGCTTC 360
Db      331 CTCTGCGACAGAGCTGGGTGCTCAAGAGATCGGCAATGAGAGTGTAGAGGCTTC 390
Qy      361 TTAGCGCAGTGTTCAGAGCCCTGCTGATCTGTATCAGTGCCTTCAAGGCTAGA 420
Db      391 TTAGCGCAGTGTTCAGAGCCCTGCTGATCTGTATCAGTGCCTTCAAGGCTAGA 450
Qy      421 AACACTACAGGAGATGAATCTTCAATAGAGAGCCATGATCTGTGTC--TTTGACATC 478
Db      451 AACACTACAGGAGATGAATCTTCAATAGAGAGCCATGATCTGTGTC--TTTGACATC 510
Qy      479 ATCAAGCCTTGTT--AGCATTTGTCAATTTTATCTTGAAGAAATCTCTGATTAAGA 536
Db      511 ATCAAGCCTTGTT--AGCATTTGTCAATTTTATCTTGAAGAAATCTCTGATTAAGA 570
Qy      537 AGATTAATTTATTA--GGTGTCTCTTCTTACCTCTGTGTGTGTGTGCGCACAGCTT 595
Db      571 AGATTAATTTATTA--GGTGTCTCTTCTTACCTCTGTGTGTGTGTGCGCACAGCTT 630
Qy      596 AGAAGGCTATTA--AGGAGAGAGAGAGTCCAAATGATGATACC--TTTATTAATTAACCATTT 654
Db      631 AGAAGGCTATTA--AGGAGAGAGAGAGTCCAAATGATGATACC--TTTATTAATTAACCATTT 690
Qy      655 CTATCAACAGGAGAGTGAAGCAGTTTC--GAGACTTTTCATGATGTTATGATGATG 712
Db      691 CTATCAACAGGAGAGTGAAGCAGTTTC--GAGACTTTTCATGATGTTATGATGATG 750
Qy      713 TTA--AAAAAGATGTTACAGTACCAATTAAGTGAAGTTTAA 754
Db      751 TTAACACAAGAGATGTTCAAGTACCAATTAAGTGAAGTTTAA 794

RESULT 6
LOCUS   B1759082 918 bp mRNA linear EST 25-SEP-2001
DEFINITION 603042879F1.NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183419 5',
RNA sequence.
ACCESSION B1759082
VERSION   B1759082.1 GI:15750660
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 918)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
COMMENT

```

```

FEATURES
source
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:5183419"
  /lab_host="DH10B"
  /note="Organ: pooled colon, kidney, stomach; Vector:
  pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
  source anonymous pool of 3 colons, age 26 yo male, 49 yo
  female, 71 yo male colon; 46 yo male kidney, and pool of 2
  stomachs, 62 yo male and 70 yo female. Library is
  oligo-dT primed and directionally cloned (EcoRV site is
  destroyed upon cloning). Average insert size 1.4 kb,
  insert size range 1-3 kb. Library is normalized and
  enriched for full-length clones and was constructed by C.
  Gruber (Invitrogen). Research Genetics tracking code
  023. Note: this is a NIH MGC Library."

BASE COUNT
269 a 196 c 247 g 206 t

ORIGIN
Query Match 87.5%; Score 659.8; DB 12; Length 918;
Best Local Similarity 96.7%; Pred. No. 6; 2e-175;
Matches 728; Conservative 0; Mismatches 17; Indels 8; Gaps 5;

1 ATGGCCGAGCTGGAGAACGAGAGAACTGCGCCCTCTCAGCTCCGGAGCTCA 60
12 ATGGCCGAGCTGGAGAACGAGAGAACTGCGCCCTCTCAGCTCCGGAGCTCA 71
61 GGCTCAAGTCGGGAGAGCCAGAGATGTTCTCCCTCAAGAGTGAAGCGGTGGCAGT 120
72 GGCTCAAGTCGGGAGAGCCAGAGATGTTCTCCCTCAAGAGTGAAGCGGTGGCAGT 131
121 TGGAGCTGGGACGTGAGTGCAGTACGTGCGCATCTGCAAGGATCCAGGTGATGCC 180
132 TGGAGCTGGGACGTGAGTGCAGTACGTGCGCATCTGCAAGGATCCAGGTGATGCC 191
181 TGTCTTAGATGTCAGCTGAAAAAACAAGAGAGCTGTGTTGTGCTGGGAGAGATG 240
192 TGTCTTAGATGTCAGCTGAAAAAACAAGAGAGCTGTGTTGTGCTGGGAGAGATG 251
241 AATCATCTCTCCACAGCTGCTGATGTCCTGTGGTGAACAGAACTGCGCCCT 300
252 AATCATCTCTCCACAGCTGCTGATGTCCTGTGGTGAACAGAACTGCGCCCT 311
301 CTCTGCGACAGAGCTGGGTGCTCAAGAGATCGGCAATGAGAGTGTAGAGGCTTC 360
312 CTCTGCGACAGAGCTGGGTGCTCAAGAGATCGGCAATGAGAGTGTAGAGGCTTC 371
361 TTAGCGCAGTGTTCAGAGCCCTGCTGATCTGTATTCAGTGCCTTCAAGGCTAGA 420
372 TTAGCGCAGTGTTCAGAGCCCTGCTGATCTGTATTCAGTGCCTTCAAGGCTAGA 431
421 AACACTACAGGAGATGAATCTTCAATAGAGAGCCATGATCTGTGTC--TTTGACATC 478
432 AACACTACAGGAGATGAATCTTCAATAGAGAGCCATGATCTGTGTC--TTTGACATC 491
479 ATCAAGCCTTGTT--AGCATTTGTCAATTTTATCTTGAAGAAATCTCTGATTAAGA 536
492 ATCAAGCCTTGTT--AGCATTTGTCAATTTTATCTTGAAGAAATCTCTGATTAAGA 551
537 AGATTAATTTATTA--AGGAGAGAGAGTCCAAATGATGATACC--TTTATTAATTAACCATTT 596
552 AGATTAATTTATTA--AGGAGAGAGAGTCCAAATGATGATACC--TTTATTAATTAACCATTT 611
597 GAAGGCTATTA--AGGAGAGAGAGTCCAAATGATGATACC--TTTATTAATTAACCATTT 655

```

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL1457 row: m column: 20
 High quality sequence stop: 732.
 Location/Qualifiers

D _b	612	GAAAGCTATATAAAGAAGGAAGGCGCCAAATTGATCAACCTTTATTAATTACCACATTC	671
O _y	656	TATATCAAAGGCAGTGGAAAGCAGTTTC - GAAGCTTTTTGATGCTTAGTGTTGATCAGT	713
D _b	672	TATATCAAAGGCAGTGGAAAGCAGTTTCAGAGAACCTTTTGATGCTTAGTGTTGATCAGT	731
O _y	714	TAA-AAAAGATGTTACAGTAAACAATAAAGNG	745
D _b	732	TAAACAAGAATGTTACAGTAAACAACATTAGAG	764

RESULT 7					
BI668735					
LOCUS	BI668735	822 bp	mRNA	linear	EST 12-SEP-2001
DEFINITION	603293174F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5312226 5',				
	mRNA sequence.				
ACCESSION	BI668735				
VERSION	BI668735.1	GI:15582968			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE	1 (pages 1 to 822)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Mikiro Palkovics, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: LNLN1791 row: d column: 19
 High quality sequence stop: 788.

FEATURES	Location/Qualifiers
source	1..822

```

219 a      182 c      219 g      202 t
BASE COUNT
BRIGIN
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5312226"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_id="NH_MGC_96"
/note="Organ: brain; Vector: pbluescript (modified
pbluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (Gcgacg
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to RQF 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NH_MGC Library."
```

Query Match	86.1%	Score 649;	DB 12;	Length 822;
Best Local Similarity	96.8%	Pred. No. 6.7e-172;		
Matches 737;	Conservative 0;	Gaps 15;	Indels 9;	Gaps 7;
1	ATGGCCGACGCGGAGACGGAGAGAAACCTGCGCCCTCGGCTCTCATCTCCGGAGCTCA	60		
29	ATGGCCGACGCGGAGACGGAGAGAAACCTGCGCCCTCGGCTCTCATCTCCGGAGCTCA	88		
61	GGCTCCCAAGTCGGAGGCGGACCAATGTTCTCCCTCAAGAAGTGGAAACGGGTGGCATG	120		
89	GGCTCCCAAGTCGGAGGCGGACCAATGTTCTCCCTCAAGAAGTGGAAACGGGTGGCATG	148		

OY	121	TGAGCTGGGACCTGGAGTGGCATCTGTCGCCATCTTCAGAGGTCACAGTGAATGGATGCC	180
Dp	149	TGAGAGCTGGGACCTGGAGTGGCATCTGTCGCCATCTTCAGAGGTCACAGTGAATGGATGCC	208
OY	181	TGCTTAGATGTCAGAGCTGAAAAACAAGAGGACTGTGTGGTGGCTGGGGAGAAATGT	240
Dp	209	TGCTTAGATGTCAGAGCTGAAAAACAAGAGGACTGTGTGGTGGCTGGGGAGAAATGT	268
OY	241	AATCATTTCTTCCACAACCTGTCATGTCCTGTGGGTGAAACAGAACAAATCGCTGCCCT	300
Dp	269	AATCATTTCTTCCACAACCTGTCATGTCCTGTGGGTGAAACAGAACAAATCGCTGCCCT	328
OY	301	CTGTGCAGACAGACTGGGTGGTCCAAAGAAATCGGACAAATGAGAGGTGATGAAAGGCTTC	360
Dp	329	CTGTGCAGACAGACTGGGTGGTCCAAAGAAATCGGACAAATGAGAGGTGATGAAAGGCTTC	388
OY	361	TTAGCGCAATGTTCCAGACCCCTGGTGGATCTTGTAATCCAGTCCCTCAGAAAGGCTAGA	420
Dp	389	TTAGCGCAATGTTCCAGACCCCTGGTGGATCTTGTAATCCAGTCCCTCAGAAAGGCTAGA	448
OY	421	ACACTACAGGGATGAATTTCTTCAATATGAGCCGATGATCTGTGGTC--TTTGAATC	478
Dp	449	ACACTACAGGGATGAATTTCTTCAATATGAGCCGATGATCTGTGGTC--TTTGAATC	508
OY	479	ATCAAAAGCCTTGTT--AGCAATTTGTCAAGTTTATCTTCAAGAAATCTCTGATTAAGA	536
Dp	509	ATCAAAAGCCTTGTT--AGCAATTTGTCAAGTTTATCTTCAAGAAATCTCTGATTAAGA	568
OY	537	AGATTAATTTAATAAAGGTGTCTTCTCACTCTGTGGTGTGTGTGGCACAACAGCTTA	596
Dp	569	AGATTAATTTAATAAAGGTGTCTTCTCACTCTGTGGTGTGTGTGGCACAACAGCTTA	628
OY	597	GAAAGTCTTAAAAA--AGAAAGAGCTCAAAATGATGATCC--TTATTAATTTCCCAATT	654
Dp	629	GAAAGTCTTAAAAA--AGAAAGAGCTCAAAATGATGATCC--TTATTAATTTCCCAATT	688
OY	655	CTATACACAGGCACTGAGACCAAGTTTCAG--ACTTTTTCGATGCTTATGGTTGATCAAT	713
Dp	689	CTATACACAGGCACTGAGACCAAGTTTCAG--ACTTTTTCGATGCTTATGGTTGATCAAT	748
OY	714	TAAAAAGAAATGTCAGTTAACAAATTAAGTCAGTTTAA	754
Dp	749	T--ACCAAGATTTACAGTTAACAAAT--AAGTCAGTTTAA	787

	RESULT	8			
	LOCUS	B1457840			
	DEFINITION	B1457840	818 bp	mRNA	linear EST 21-AUG-2006 603198212F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5277723 5' , mRNA sequence.
	ACCESSION	B1457840			
	VERSION	B1457840.1	GI:15248496		
	KEYWORDS	EST.			
	SOURCE	Homo sapiens (human)			
	ORGANISM	Homo sapiens			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 818) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished			
	JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgabbs-re@mail.nih.gov Tissue Procurement: Miklos Pal Kovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshiyuki and Piero Carninci (RIKEN) CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLt) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLt at: http://image.llnl.gov Plate: LLMML1701 Row: g Column: 04 High quality sequence stop: 780.			
	COMMENT				
	REFERENCE				
	AUTHORS				
	TITLE				

QY 241 AATCATCTCTTCCACACTGCTGATGTCTCTGTGGTGAACAGAACATGCTGCTCCT 300
 DB 266 AATCATCTCTTCCACACTGCTGATGTCTCTGTGGTGAACAGAACATGCTGCTCCT 325
 QY 301 CTCTGCGACAGAGACTGGGTGGTCCAAAGATCGGCAATGAGAGGTGTAAGAGCTTC 360
 DB 326 CTCTGCGACAGAGACTGGGTGGTCCAAAGATCGGCAATGAGAGGTGTAAGAGCTTC 385
 QY 361 TTAGGCGAGTGTTCAGAGCCCTGTGTGATCTTGAATCAGTGCCTTCAAAAGCTTGA 420
 DB 386 TTAGGCGAGTGTTCAGAGCCCTGTGTGATCTTGAATCAGTGCCTTCAAAAGCTTGA 445
 QY 421 AACATACAGGAGATGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGACATC 478
 DB 446 AACATACAGGAGATGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGACATC 505
 QY 479 AACAAAGCCTTGCTT--AGCATTTGTGATTTTATCTTCAAGAAATCTCTGTGATTAAGA 536
 DB 506 ATCAAAAGCCTTGCTT--AGCATTTGTGATTTTATCTTCAAGAAATCTCTGTGATTAAGA 565
 QY 537 AGATATATTTATTA--AGTGTGCTCTTCTTACCTCTGTGTGTGTGTGCGCACAGCTT 595
 DB 566 AGATATATTTATTA--AGTGTGCTCTTCTTACCTCTGTGTGTGTGTGCGCACAGCTT 625
 QY 596 AGAAGTGTATTAATAAAGAAAGAGTCCAAATGATGATCACC-TTATATTTATCCATTT 654
 DB 626 AGAAGTGTATTAATAAAGAAAGAGTCCAAATGATGATCACC-TTATATTTATCCATTT 685
 QY 655 CTATACACAGGAGAGTGAAGAGCTTTCGAG---ACTTTTGCATGCTTATGTTGATC 710
 DB 686 CTATACACAGGAGAGTGAAGAGCTTTCGAG---ACTTTTGCATGCTTATGTTGATC 745
 QY 711 AGTTAAAAAAGATGTTACAGTACCAATTAAGTGCATTTAA 754
 DB 746 AGTTAAAAAAGATGTTACAGTACCAATTAAGTGCATTTAA 789

RESULT 10
 B1601855 856 bp mRNA linear EST 07-SHP-2001
 LOCUS 603244729F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287014 5',
 DEFINITION mRNA sequence.
 ACCESSION B1601855
 VERSION B1601855.1 GI:15494794
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 856)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Niklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshitaki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM1725 row: j column: 07
 High quality sequence stop: 820.
 Location/Qualifiers
 1. 856
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5287014"
 /tissue_type="hypothalamus"

FEATURES
 source

/lab host="DH10B"
 /clone_lib="NIH_MGC_96"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTTTV-3',
 size selected for average insert size 2.3 kb and
 normalized to 10^5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 227 a 186 c 227 g 216 t
 ORIGIN
 Query Match 85.3%; Score 643; DB 12; Length 856;
 Best Local Similarity 95.3%; Pred. No. 3.4e-170;
 Matches 729; Conservative 0; Mismatches 25; Indels 11; Gaps 6;

QY 1 ATGGCCGACGCGGAAAGACGAGAGAAACCTGGCCCTGCTCACTCCGGAGCTCA 60
 DB 17 ATGGCCGACGCGGAAAGACGAGAGAAACCTGGCCCTGCTCACTCCGGAGCTCA 76
 QY 61 GGCCTCAAGTCGGAGAGCGCAAGATGTTCTCCCTCAAGAGAGTGAACGCGGTGCCATG 120
 DB 77 GGCCTCAAGTCGGAGAGCGCAAGATGTTCTCCCTCAAGAGAGTGAACGCGGTGCCATG 136
 QY 121 TGAAGCTGGGACGTGAGATGCGATGCGCCATCTGACAGGTTCAGGTATGATGCC 180
 DB 137 TGAAGCTGGGACGTGAGATGCGATGCGCCATCTGACAGGTTCAGGTATGATGCC 196
 QY 181 TGTCTTAGATGTCAGAGCTGAAACAAACAAAGAGAGCTGTGTTGGTGGGAGAAATGT 240
 DB 197 TGTCTTAGATGTCAGAGCTGAAACAAACAAAGAGAGCTGTGTTGGTGGGAGAAATGT 256
 QY 241 AATCATCTCTTCCACACTGCTGATGTCTCTGTGGTGAACAGAACATGCTGCTCCT 300
 DB 257 AATCATCTCTTCCACACTGCTGATGTCTCTGTGGTGAACAGAACATGCTGCTCCT 316
 QY 301 CTCTGCGACAGAGACTGGGTGGTCCAAAGATCGGCAATGAGAGGTGTAAGAGCTTC 360
 DB 317 CTCTGCGACAGAGACTGGGTGGTCCAAAGATCGGCAATGAGAGGTGTAAGAGCTTC 376
 QY 361 TTAGGCGAGTGTTCAGAGCCCTGTGTGATCTTGAATCAGTGCCTTCAAAAGCTTGA 420
 DB 377 TTAGGCGAGTGTTCAGAGCCCTGTGTGATCTTGAATCAGTGCCTTCAAAAGCTTGA 436
 QY 421 AACATACAGGAGATGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGACATC 477
 DB 437 AACATACAGGAGATGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGACATC 496
 QY 478 CATCAAGCCTTGCTT--AGCATTTGTGATTTTATCTTCAAGAAATCTCTGTGATTAAG 535
 DB 497 CATCAAGCCTTGCTT--AGCATTTGTGATTTTATCTTCAAGAAATCTCTGTGATTAAG 556
 QY 536 AAGATATTTATTA--AGTGTGCTCTTCTTCACTCTGTGTGTGTGTGCGCACAGACT 594
 DB 557 AAGATATTTATTA--AGTGTGCTCTTCTTCACTCTGTGTGTGTGTGCGCACAGACT 616
 QY 595 TAGAAGTCTTATTAATAAAGAAAGAGCTCCAAATGATGATCACC-TTATATTTATCCATT 653
 DB 617 TAGAAGTCTTATTAATAAAGAAAGAGCTCCAAATGATGATCACC-TTATATTTATCCATT 676
 QY 654 TCTATACACAGGAGAGTGAAGAGCTTTC--GAGACTTTTTCATGCTTATG--TTGAT 709
 DB 677 TCTATACACAGGAGAGTGAAGAGCTTTC--GAGACTTTTTCATGCTTATG--TTGAT 736
 QY 710 CAGTTAAAAAAGATGTTACAGTAACTTAAAGTGCAGTTAA 754
 DB 737 CAGTTAAAAAAGATGTTACAGTAACTTAAAGTGCAGTTAA 781

RESULT 11
 BG766992

LOCUS BG766992 754 bp mRNA linear EST 15-MAY-2001
 DEFINITION 60274051F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4870251 5',
 mRNA sequence.
 ACCESSION BG766992
 VERSION BG766992.1 GI:14077645
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DRP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLCM1743 row: e column: 04
 High quality sequence stop: 750.
 Location/Qualifiers
 1..754
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4870251"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_49"
 /note="Organ: Skin; Vector: pOT7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

BASE COUNT 205 a 160 c 197 g 192 t

ORIGIN

Query Match 84.9%; Score 640.2; DB 12; Length 754;
 Best Local Similarity 96.9%; Pred. No. 2e-169;
 Matches 728; Conservative 0; Mismatches 13; Indels 10; Gaps 7;

13 GAAGACGAGAGAAACCTGCGCCCTGCTCACTCCGGAGCTCAGGCTTCAAGTCG 72
 1 GAAGACGAGAGAAACCTGCGCCCTGCTCACTCCGGAGCTCAGGCTTCAAGTCG 60
 73 GGAGCGCAAGAAGTGTCTCCCTCAAGAAGTGAAGCGGTGGCCATGTGGAGCTGGAC 132
 61 GGAGCGCAAGAAGTGTCTCCCTCAAGAAGTGAAGCGGTGGCCATGTGGAGCTGGAC 120
 133 GTGAGTGGCATCTGCGCCATCTGCAAGGCTCAGGTATGATGATCTTATAGT 192
 121 GTGAGTGGCATCTGCGCCATCTGCAAGGCTCAGGTATGATGATCTTATAGT 180
 193 CAAGCTGAAAACAAACAAAGAGAGCTGTGTGTGTGTGGGAGATGTATCATCTTC 252
 181 CAAGCTGAAAACAAACAAAGAGAGCTGTGTGTGTGTGGGAGATGTATCATCTTC 240
 253 CAAGCTGCGCATGTCCCTGTGGTGAAGAGAAACAAATCGTGCCTCTGCGCAGCG 312
 241 CACAACTGCTGATGTCCCTGTGGGAGAAACAAATCGTGCCTCTGCGCAGCG 300
 313 GACTGGTGTCTCAAAAGATCGGCAATGAGAGTGTAGAGGCTTCTTACGCGAGTTG 372
 301 GACTGGTGTCTCAAAAGATCGGCAATGAGAGTGTAGAGGCTTCTTACGCGAGTTG 360

373 TTCAGAGCCCTGGTGTCTTGTATTCAGTGCCTTCAAAAGGCTAGAACACTACAGGG 432
 361 TTCAGAGCCCTGGTGTCTTGTATTCAGTGCCTTCAAAAGGCTAGAACACTACAGGG 420
 433 ATGAATCTTCAAAATAGAGCCGATGATCTGTGTC--TTTGACTCATCAAGCCTTG 490
 422 ATGAATCTTCAAAATAGAGCCGATGATCTGTGTC--TTTGACTCATCAAGCCTTG 480
 491 GTT--AGCATTTGTCTGATTTATCTTCAAGAAATCTCTGTGATTA--GAAGATATTTAT 547
 481 GTTACCATTTTGTCTGATTTATCTTCAAGAAATCTCTGATTA--GAAGATATTTAT 540
 548 TAA--GGGTGCTCTCTACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 606
 541 TAAAGGTGTCT 600
 607 AAAAAAGAAAGAGCTCAAAATGATCAAC--TTATATTTACCCATTTCTATCAACAG 665
 601 AAAAAAGAAAGAGCTCAAAATGATCAACCTTTATATTTACCCATTTCTATCAACAG 660
 666 GCAGTGGAAGCAGTTTC--GAGACTTTTTCATGCTTATGTTGATCAGTTAAAAAGAA 723
 661 GCAGTGGAAGCAGTTTC--GAGACTTTTTCATGCTTATGTTGATCAGTTAAAAAGAA 720
 724 TGTTCAGTAAACAAATTAAGTCAGTTTAA 754
 721 TGTTCAGTAAACAAAT--AAGTCAGTTTAA 750

RESULT 12
 BMS56683
 LOCUS BMS56683 1016 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGENCOURT_6540688 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5737890
 5', mRNA sequence.
 ACCESSION BMS56683
 VERSION BMS56683.1 GI:18798111
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLAH12748 row: 1 column: 19
 High quality sequence stop: 620.
 Location/Qualifiers
 1..1016
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5737890"
 /tissue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_88"
 /note="Organ: small intestine; Vector: pCMV-SPORT6;
 Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
 oligo-dT primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 271 a 231 c 266 g 246 t

ORIGIN

Query Match 84.3%; Score 635.4; DB 12; Length 1016;
 Best Local Similarity 95.5%; Pred. No. 5e-168;
 Matches 721; Conservative 0; Mismatches 21; Indels 13; Gaps 6;

1 ATGGCCGACGCTGGAGAGAGAGAACTGCGCCCTGCGCTCTACATCCGGAGACTCA 60
 |||||
 52 ATGGCCGACGCTGGAGAGAGAGAACTGCGCCCTGCGCTCTACATCCGGAGACTCA 111
 |||||
 61 GGCTCCAGATCGGAGGCGCAAGATGTTCTCCCTCAAGAGTGAACGCGGTGCGCATG 120
 |||||
 112 GGCTCCAGATCGGAGGCGCAAGATGTTCTCCCTCAAGAGTGAACGCGGTGCGCATG 171
 |||||
 121 TGGAGCTGGAGAGTGGAGTGCATATCGTCGCCATCTGCAAGGCTCCAGTGAATGATGCC 180
 |||||
 172 TGGAGCTGGAGAGTGGAGTGCATATCGTCGCCATCTGCAAGGCTCCAGTGAATGATGCC 231
 |||||
 181 TGTCTAGATGTCAGAGCTGGAACCAAGAGAGAGAGTGTGTTGCTGTGGGAGAAATGT 240
 |||||
 232 TGTCTAGATGTCAGAGCTGGAACCAAGAGAGAGTGTGTTGCTGTGGGAGAAATGT 291
 |||||
 241 AATCATTCCTTCCACAACCTGCTGATGTCCTGCTGGGTGAACAGAACATCGTGCCT 300
 |||||
 252 AATCATTCCTTCCACAACCTGCTGATGTCCTGCTGGGTGAACAGAACATCGTGCCT 351
 |||||
 301 CTCTGCGACGAGAGCTGGGTGCTCCAAAGATCGCAATGAGAGTGTAGAGGCTTC 360
 |||||
 352 CTCTGCGACGAGAGCTGGGTGCTCCAAAGATCGCAATGAGAGTGTAGAGGCTTC 411
 |||||
 361 TTAGGCGAGTGTTCAGAGCCCTGGTGGATCTTTGATCAGTGGCCCTCAAGGCTTGA 420
 |||||
 412 TTAGGCGAGTGTTCAGAGCCCTGGTGGATCTTTGATCAGTGGCCCTCAAGGCTTGA 471
 |||||
 421 ACACCTACAGGAGATGATTTCTTCAATAGAGCCGATGATGATGCTG- -TTTGGACTC 478
 |||||
 472 ACACCTACAGGAGATGATTTCTTCAATAGAGCCGATGATGATGCTG- -TTTGGACTC 531
 |||||
 479 ATCAAGAGCTTGGT- -ACATTTGTCACTTTTATCTTCAAGATTTCTGTGATTAAGA 536
 |||||
 532 ATCAAGAGCTTGGT- -ACATTTGTCACTTTTATCTTCAAGATTTCTGTGATTAAGA 591
 |||||
 537 AGATATTTATTAAGAGTGGTCTTCTCAACCTCTGT- -GGTGTGTGGGCGACAGAGCT 595
 |||||
 552 AGATATTTATTAAGAGTGGTCTTCTCAACCTCTGTGTGTGTGTGTGTGTGTGTGTGT 651
 |||||
 556 AGAAGTGTCTATTAAGAGAGAGAGAGCTCCAAATTGATCACC- -TTATATTTATCCATT 654
 |||||
 652 AGAAGTGTCTATTAAGAGAGAGAGAGAGCTCCAAATTGATCACC- -TTATATTTATCCATT 711
 |||||
 655 CTATACCAAG 708
 |||||
 712 CTATACCAAG 771
 |||||
 709 TCAG- -TTAAAAAGAGATGTTACAGTAAACAAATAA 742
 |||||
 772 TCAGTTTAAAAAGAGATGTTACAGTAAACAAATAA 806
 |||||

RESULT 13
 BM704660 706 bp mRNA linear EST 28-FEB-2002
 LOCUS UT-E-C11-af0-o-15-0-UI.r2 UT-E-C11 Homo sapiens cDNA clone
 DEFINITION UT-E-C11-af0-o-15-0-UI 5', mRNA sequence.
 ACCESSION BM704660
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 706)
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene

JOURNAL discovery
 MEDLINE Genome Res. 6 (9), 791-806 (1996)
 PUBMED 97044477
 COMMENT 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..706
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UT-E-C11-af0-o-15-0-UI"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UT-E-C11"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UT-E-C11 is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Ronaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACTTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 182 a 153 c 191 g 180 t
 ORIGIN

Query Match 84.1%; Score 634.4; DB 12; Length 706;
 Best Local Similarity 97.5%; Pred. No. 8.3e-168;
 Matches 688; Conservative 0; Mismatches 11; Indels 7; Gaps 4;

5 CCGACGTGGAAGAGAGAGAGAACTGCGCCCTGCGCTCTACATCCGGAGACTCAGGCT 64
 |||||
 1 CCGACGTGGAAGAGAGAGAGAACTGCGCCCTGCGCTCTACATCCGGAGACTCAGGCT 60
 |||||
 65 CCAAGTGGGAGGCGACAGATGTTCTCCCTCAAGAGTGAACGCGGTGCGCATGTCGA 124
 |||||
 61 CCAAGTGGGAGGCGACAGATGTTCTCCCTCAAGAGTGAACGCGGTGCGCATGTCGA 120
 |||||
 125 GCTGGAGCGTGGAGTGGATCGGCGGCATCTGCAAGGCTCAGGTATGATGATGCTGTC 184
 |||||
 121 GCTGGAGCGTGGAGTGGATCGGCGGCATCTGCAAGGCTCAGGTATGATGATGCTGTC 180
 |||||
 185 TTAGATGTCAAGCTGAAGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 244
 |||||
 181 TTAGATGTCAAGCTGAAGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 |||||
 245 ATTCTTCCACAATGCTGCTGATGTCCTGCTGGGTGAACAGAAACAAATGCTGCTCTCT 304
 |||||
 241 ATTCTTCCACAATGCTGCTGATGTCCTGCTGGGTGAACAGAAACAAATGCTGCTCTCT 300
 |||||
 305 GCCAGCGAGAGCTGGGTGTCCAAGAGATCGGCAATGAGAGTGTAGAGAGCTTCTTAG 364
 |||||

DB	301	GCACGACGAGCTGGGTGGTCCAAAGAAATCGGCAAAATGAGAGTGGTTAGAAAGCTTCTTAG	360
QY	365	CGCAGTTGTTGACAGCCCTGGTGGATCTTTGATATCATGCTGCCCTTACAAAGCTTGAACAC	424
Db	361	CGCAGTTGTTGACAGCCCTGGTGGATCTTTGATATCATGCTGCCCTTACAAAGCTTGAACAC	420
QY	425	TACAGGGGATGAATCTTCAATAGAGCCGATGATCTGGTGC--TTTGACATCA	482
Db	421	TACAGGGGATGAATCTTCAATAGAGCCGATGATCTGGTGC--TTTGACATCA	480
QY	483	AAGCCTTGTT--AGCATTTGTCAGTTTATCTTACAGAAATCTCTGATTAAGAAAGAT	540
Db	481	AAGCCTTGTTGATGATTTTGTGCACTTTATCTTACAGAAATCTCTGATTAAGAAAGAT	540
QY	541	AATTTATTAAGGTGGTCTCTTCTTCACTCTGTGGTGTGTCGCGCACAGCTTAGAG	600
Db	541	AATTTATTAAGGTGGTCTCTTCTTCACTCTGTGGTGTGTCGCGCACAGCTTAGAG	600
QY	601	TGCTATTAAGGAAAGAAAGAGCTCCAAATTTGAATCAC--TTATATTTACCATTTCTATA	659
Db	601	TGCTATTAAGGAAAGAAAGAGCTCCAAATTTGAATCAC--TTATATTTACCATTTCTATA	660
QY	660	CACAGCGAGTGGAGCAGATTTC--GAGACTTTTCGATGGCTTAG	703
Db	661	CACAGCGAGTGGAGCAGATTTCGAGAGAACTTTTTCATGCTTAG	706
RESULT 14			
LOCUS	B1226556	724 bp	mRNA linear EST 11-JUL-2001
DEFINITION	602951527F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:5095687 5',		
ACCESSION	B1226556		
VERSION	B1226556.1	GI:14680000	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 724)		
JOURNAL	NIH-MGC http://mgi.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cga@bbs-rcmail.nih.gov		
	Tissue Procurement: Louis M. Staudt, M.D., Ph.D.		
	cDNA Library Preparation: Ling Hong/Rubin Laboratory		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.llnl.gov		
	Plate: L1CML1866 row: f column: 08		
	High quality sequence stop: 720.		
FEATURES	Location/Qualifiers		
source	1..724		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone_image="5095687"		
	/tissue_type="Burkitt lymphoma"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH MGC 8"		
	/note="Organ: lymph; Vector: pOT7; Site 1: XhoI; Site 2:		
	EcoRI; cDNA made by oligo-dt priming. Directionally		
	cloned into EcoRI/XhoI sites using the following 5'		
	adaptor: GGCACGAG(G). Size-selected >500bp for average		
	insert size 1.8kb. Library constructed by Ling Hong in		
	the laboratory of Gerald M. Rubin (University of		
	California, Berkeley) using ZAP-cDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies)."		
BASE COUNT	195 a 153 c 189 g 187 t		
ORIGIN			

Query Match	84.1%;	Score 634.4;	DB 12;	Length 724;
Best Local Similarity	96.6%;	Pred. No. 8.4e-168;		
Matches 705; Conservative	0;	Mismatches 11;	Indels 14;	Gaps 5;

[illegible]


```

QY 210 AGAGACTGTGTGTGGTGGGAGATATATCATTCCTTCCACAACAGTGTGATGTC 269
    |||||
Db 181 AGAGACTGTGTGTGTGGTGGGAGATATATCATTCCTTCCACAACAGTGTGATGTC 240
    |||||
QY 270 CCTGTGGTGAACAGAAACAACTGCTGCTCTCTCCAGACAGAGTGGTGGTCCAAAG 329
    |||||
Db 241 CCTGTGGTGAACAGAAACAACTGCTGCTCTCTCCAGACAGAGTGGTGGTCCAAAG 300
    |||||
QY 330 AATCGGCAATGAGTGTGTAGAGGCTTCTTAGGCGAGTGTTCAGAGCCCTGTGTGA 389
    |||||
Db 301 AATCGGCAATGAGTGTGTAGAGGCTTCTTAGGCGAGTGTTCAGAGCCCTGTGTGA 360
    |||||
QY 390 TCTTGAATTCAGTGGCCCTACAAAGGCTAGAACACTACAGGGAGTGAATTCCTCAATAG 449
    |||||
Db 361 TCTTGAATTCAGTGGCCCTACAAAGGCTAGAACACTACAGGGAGTGAATTCCTCAATAG 420
    |||||
QY 450 GAGCCGATGATCTGTGTGTC--TTTGGACTCATCAAGCCTTGG--TTAGCATTTGTCAAT 506
    |||||
Db 421 GAGCCGATGATCTGTGTGTC--TTTGGACTCATCAAGCCTTGG--TTAGCATTTGTCAAT 480
    |||||
QY 507 TTTATCTTCAGAAATCTCTGTGATTAAGAGTAAATTTATTAAGTGTCTCTTCTTAC 566
    |||||
Db 481 TTTATCTTCAGAAATCTCTGTGATTAAGAGTAAATTTATTAAGTGTCTCTTCTTAC 540
    |||||
QY 567 CTCTGTGTGTGTGTGTGGCGACACAGCTTGAAGTGTATTAAGAAAGAAAGAGCTCCAA 626
    |||||
Db 541 CTCTGTGTGTGTGTGTGGCGACACAGCTTGAAGTGTATTAAGAAAGAAAGAGCTCCAA 600
    |||||
QY 627 ATTGAATCAAC--TTATATTTTACCAATTTCTATACACA--GGCAGTGAAGCAGTTTC-- 682
    |||||
Db 601 ATTGAATCAACCTTTATATTTATTCATCCATTTCTATACACAAGGGCAGTGAAGCAGTTTCAG 660
    |||||
QY 683 GAGACTTTTTCAGTCTGTATGTTGATCACTTAAAGAAAGATTTTACAGTACAA 737
    |||||
Db 661 AGAATCTTTTTCAGTCTGTATGTTGATCACTT--AAAAGAAATTTTACCGTAAAAA 714
    |||||

RESULT 17
LOCUS BG037017 893 bp mRNA linear EST 24-JAN-2001
DEFINITION BG037017.1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4374483 5',
            mRNA sequence.
ACCESSION BG037017
VERSION BG037017.1 GI:12432822
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 893)
          NIH-MGC http://mgs.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished
          Contact: Robert Strausberg, Ph.D.
          Email: cgaabs@remail.nih.gov
          Tissue Procurement: Niklos Palkovits, M.D., Ph.D.
          cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
          Toshinuki and Piero Carninci (RIKEN)
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LIML at:
          http://image.jiml.gov
          Plate: LHAM10038 row: d column: 04
          High quality sequence stop: 741.
          Location/Qualifiers
            1..893
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:4374483"
              /issue_type="hypothalamus"

```

```

/!ab host="DH10B"
/!note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT
  241 a 202 c 253 g 197 t
ORIGIN
Query Match      82.7%; Score 623.6; DB 10; Length 893;
Best Local Similarity 97.1%; Pred. No. 16-164;
Matches 678; Conservative 0; Mismatches 14; Indels 6; Gaps 4;

QY 1 ATGGCCGACGTGGAAGACGAGAGAAACCTGCGCTGAGCTCTCACTCCGAGACTCA 60
    |||||
Db 29 ATGGCCGACGTGGAAGACGAGAGAAACCTGCGCTGAGCTCTCACTCCGAGACTCA 88
    |||||
QY 61 GGCTCCAACTCGGAGGCGCAAGATGTTCTCTCTCAAGAACTGGAACCGGCTGGCATG 120
    |||||
Db 89 GGCTCCAACTCGGAGGCGCAAGATGTTCTCTCTCAAGAACTGGAACCGGCTGGCATG 148
    |||||
QY 121 TGGAGCTGGAGAGTGGAGTGCATACGTGCGGCATCTGAGAGGTCCAGTATGATGTC 180
    |||||
Db 149 TGGAGCTGGAGAGTGGAGTGCATACGTGCGGCATCTGAGAGGTCCAGTATGATGTC 208
    |||||
QY 181 TGTCTTAGATGTCAGAGCTGAAACAAACAAGAGACTGTGTGGTGTGGGAGAAATGT 240
    |||||
Db 209 TGTCTTAGATGTCAGAGCTGAAACAAACAAGAGACTGTGTGGTGTGGGAGAAATGT 268
    |||||
QY 241 AATCAATCTCTTCCAACTGCTGATGTCCTGTGGTGAACAGAAACATCTGCTCCCT 300
    |||||
Db 269 AATCAATCTCTTCCAACTGCTGATGTCCTGTGGTGAACAGAAACATCTGCTCCCT 328
    |||||
QY 301 CTCTGCGACAGAGACTGGGTGCTCCAAAGAAATCGGCAATGAGTGTGTTAAGGCTTC 360
    |||||
Db 329 CTCTGCGACAGAGACTGGGTGCTCCAAAGAAATCGGCAATGAGTGTGTTAAGGCTTC 388
    |||||
QY 361 TTAGGCGAGTGTTCAGAGCCCTGTGGATCTGTATATCACTGACCTTCAAGAGCTAGA 420
    |||||
Db 389 TTAGGCGAGTGTTCAGAGCCCTGTGGATCTGTATATCACTGACCTTCAAGAGCTAGA 448
    |||||
QY 421 ACACCTACAGGGGATGAATTTCTTCAATAGAGCCGATGATCTGTGTC--TTTGGACTC 478
    |||||
Db 449 ACACCTACAGGGGATGAATTTCTTCAATAGAGCCGATGATCTGTGTC--TTTGGACTC 508
    |||||
QY 479 ATCAAAAGCTTGGTT--AGCATTTGTCACTTTTATCTTCAAGAAATTTCTGTGATTAAGA 536
    |||||
Db 509 ATCAAAAGCTTGGTTTATGATTTGTCACTTTTATCTTCAAGAAATTTCTGTGATTAAGA 568
    |||||
QY 537 AATATATTATTA--AGGAGTCTTCTTCACTCTGTGTGTGTGTGGCGACACAGCTT 595
    |||||
Db 569 AATATATTATTAACAGAGTGTCTTCTTCACTCTGTGTGTGTGTGGCGACACAGCTT 628
    |||||
QY 596 AGAAGTGTATTAAGAAAGAAAGAGCTCCAAATTTGATCAAC--TTATATTTATTTACCATTT 654
    |||||
Db 629 AGAAGTGTATTAAGAAAGAAAGAGCTCCAAATTTGATCAACCTTTATATTTATTTACCATTT 688
    |||||
QY 655 CTATACACAGGCACTGGAAGCACTTTTGAAGCTTTT 692
    |||||
Db 689 CTATACACAGGCACTGGAAGCACTTTTGAAGCACTTTT 726
    |||||

RESULT 18
LOCUS BX387707/c 1201 bp mRNA linear EST 08-MAY-2003
DEFINITION BX387707 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS01037YPI3 3-PRIME, mRNA sequence.
ACCESSION BX387707

```

```

VERSION      EX387707.1  GI:30463313
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        1 (bases 1 to 1201)
JOURNAL      Li W.B., Gruber C., Jesse J. and Polyes D.
COMMENT      Full-length cDNA libraries and normalization
              Unpublished
              Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: sequef@genoscope.cns.fr Web: www.genoscope.cns.fr
              was normalized. Library was constructed by Life Technologies, a
              division of Invitrogen. This sequence belongs to sequence cluster
              2013.f. For more information about this cluster, see
              http://www.genoscope.cns.fr/
              cgi-bin/cluster.cgi?seq=CS0A1037CH07NP1&cluster=2013.f. Contact :
              Feng Liang Email: fliang@lifetech.com URL:
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue Genoscope sequence ID : CS0A1037CH07NP1.
              Location/Qualifiers
FEATURES
  source
    1..1201
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CS0D1037YPL3"
    /tissue_type="PLACENTA COT 25-NORMALIZED"
    /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
    /note="1st strand cDNA was primed with a NotI-cloigo (dr)
    primer: Five prime end enriched, double-strand cDNA was
    digested with Not I and cloned into the Not I and EcoR V
    sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT
  283 a 249 c 234 g 299 t 136 others
Query Match      82.1%; Score 619.4; DB 13; Length 1201;
Best Local Similarity 89.6%; Pred. No. 1.8e-163;
Matches 678; Conservative 6; Mismatches 66; Indels 7; Gaps 4;
QY 1 ATGCGCGAGTGGAAAGCGAGAGGAAACCTGCGCCCTCTCACTCCCGGAGCTCA 60
DB 757 ATGGCCMACGTGAGARACGAGAGAAACCTGCGCCCTCTCACTCCCGGAGCTCA 698
QY 61 GGCTCCAGTGGGAGGCGCAAGATGTTCTCCCTCAAGAAATGGAAGCGGTCCTATG 120
DB 697 GGCTCCAGTGGGAGGCGCAAGATGTTCTCCCTCAAGAAATGGAAGCGGTCCTATG 638
QY 121 TGGAGCTGGGAGCGTGGAGTGCATGCGCGCATCTGAGAGGTCGAGGTGATGATGCC 180
DB 637 TGGGCGTGGMACGTGAGTGCATGCGCGCATCTGAGAGGTCGAGGTGATGATGCC 578
QY 181 TGTCTTGAATGTCAAGCTGAAACAAACAGAGAGCTGTGTGTTGTCGCGGAATGT 240
DB 577 TGTCTTGAATGTCAAGCTGAAACAAACAGAGAGTGTGTGTTGTCGCGGAATGT 518
QY 241 AATCATTCCTCCCAACCTGTCGATGTCCTGTCGGTGGGAAACAGAAATTCGTCGCTC 300
DB 517 AATCATTCCTCCCAACCTGTCGATGTCCTGTCGGTGGGAAACAGAAATTCGTCGCTC 458
QY 301 CTCTCCAGCAGAGAGCTGGGTGGTCCAAAGAAATCGGCAATGAGTGGTTGAAGGCTTC 360
DB 457 NTCTGCCAGCAGAGAGCTGGGTGGTCCAAAGAAATCGGCAATGAGTGGTTGAAGGCTTC 398
QY 361 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTGAATCCAGTGCCTCAACAAGCTTGA 420
DB 397 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTGAATCCAGTGCCTCAACAAGCTTGA 338
QY 421 ACACACAGAGGGGAGTAATTTCTCAATAGAGCGCATGAGATCTGTGTC--TTTGGACTC 478
DB 337 AACTACAGAGGGGAGTAATTTCTCAATAGAGCGCATGAGATCTGTGTC--TTTGGNATTC 278

```

```

QY 479 ATCAAGCCTTGTT--AGCATTTGTCAGTTTATCTTCAAGAAATCTCTGTGATTAGA 536
DB 277 ATNAAAGCCTTGTTTACATTTTNTGAGTTTATCTTCAAGAAATTTCTCTCGATTAGA 218
QY 537 AGATATTTATTAAGAGGTCCTTCTACCTCTGTCGTGTGTGTCGCGACACAGCTTA 596
DB 217 AGATATTTATTAAGAGGTCCTTCTACCTCTGTCGTGTGTGTCGCGACACAGCTTA 158
QY 597 GAAGTGTATTAAGAAAGAGAGCTCCAAATGAATCAC--TTATATTTTACCATTTC 655
DB 157 GAAGTGTATTAAGAAAGAGAGTCTCCAAATGAATCACCTTTATATTTACCATTTC 98
QY 656 TATACAAAGCAGTGGAGAGAGTTTC--GAGACTTTTTCATGCTTATGTTGATCAT 713
DB 97 TATCTACAGCAGTGGAGAGAGTTTCAGAGAACTTTTTCATGCTTATGTTGATCAT 38
QY 714 TAAAAAGATGTTTACAGTACATTAATTAAGTGCAGTT 750
DB 37 TAAAAAGATGTTTACAGTACATTAATTAAGTGCAGTT 1

```

```

RESULT 19
LOCUS      B1858307 797 bp mRNA linear EST 10-OCT-2001
DEFINITION 603384080P1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5393131 5',
            mRNA sequence.
ACCESSION  B1858307
VERSION    B1858307.1 GI:15999054
KEYWORDS  EST.
SOURCE     Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL    National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT    Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: gga@lelmail.nih.gov
            Tissue Procurement: DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: L1AM12001 row: 0 column: 20
            High quality sequence, stop: 718.
            Location/Qualifiers
            1..797
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5393131"
            /tissue_type="mammary adenocarcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_87"
            /note="Organ: Breast; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
            Average insert size 1.383 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC library."
BASE COUNT  229 a 166 c 204 g 198 t
ORIGIN

```

```

Query Match      81.7%; Score 616.2; DB 12; Length 797;
Best Local Similarity 95.0%; Pred. No. 1.2e-162;
Matches 725; Conservative 0; Mismatches 23; Indels 15; Gaps 8;
QY 4 GCGAGCTGGAAGACGAGAGGAAACCTGCGCCCTGCTCACTCCGAGACTCAGGC 63
DB 1 GCGAGCTGGAAGACGAGAGGAAACCTGCGCCCTGCTCACTCCGAGACTCAGGC 60

```


QY 64 TCCAGTGGGAGCGCAAGATGTTCTCCCTCAAGAGTGAACGGGTGGCCATGTG 123
 DB 61 TCCAGTGGGAGCGCGCAAGATGTTCTCCCTCAAGAGTGAACGGGTGGCCATGTG 120
 QY 124 AGCTGGGACGTGAGTGGCATACGTGGCCATCTGCAGGGTCCAGTATGATGCTGT 183
 DB 121 AGCTGGGACGTGAGTGGCATACGTGGCCATCTGCAGGGTCCAGTATGATGCTGT 180
 QY 184 CTAGATGTCACCTGAAAAAACAAGAGACCTGTGTGTGCTGGGGAGATGAT 243
 DB 181 CTAGATGTCACCTGAAAAAACAAGAGACCTGTGTGTGCTGGGGAGATGAT 240
 QY 244 CATTCCTCCACAACCTGTCATGTCCTGTGGGTGAACAACAATCGCTCCCTTC 303
 DB 241 CATTCCTCCACAACCTGTCATGTCCTGTGGGTGAACAACAATCGCTCCCTTC 300
 QY 304 TGGCAGACGAGCTGGGTGTGTCACAAGATGCGCAATGAGAGTGTAGAGGCTTCTTA 363
 DB 301 TGGCAGACGAGCTGGGTGTGTCACAAGATGCGCAATGAGAGTGTAGAGGCTTCTTA 360
 QY 364 GGGCAGTGTGTCAGAGCCCTGGTGGATCTTGTATCAAGTCCCTCAAAAGGCTAGAAC 423
 DB 361 GGGCAGTGTGTCAGAGCCCTGGTGGATCTTGTATCAAGTCCCTCAAAAGGCTAGAAC 420
 QY 424 CTACAGGAGTGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGGACTCATC 481
 DB 421 CTACAGGAGTGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGGACTCATC 480
 QY 482 AAGGCTTGCTT--AGCATTTGTCAGTTTATCTTCAAGAAATCTCTGTGATTAAGAAG 539
 DB 481 AAGGCTTGCTT--AGCATTTGTCAGTTTATCTTCAAGAAATCTCTGTGATTAAGAAG 540
 QY 540 TAATTAAT--AAGGAGTGTCTCTCAACCTGTGGTGTGTGCGGACACAGCTTAA 598
 DB 541 TAATTAAT--AAGGAGTGTGTCTCTCAACCTGTGGTGTGTGCGGACACAGCTTAA 600
 QY 599 AGTGTATAAAAAAGGAAA--GAGCTCCAATTAATCAAC--TTATATTTAACCATTCT 656
 DB 601 AGTGTATAAAAAAGGAAA--GAGCTCCAATTAATCAAC--TTATATTTAACCATTCT 660
 QY 657 ATACAAAGGACGATGAGAGAGTTTC--GAGACTTTTTCAGTCTTATGTTATCATGTT 714
 DB 661 AT--ACAGGACGATGAGAGAGTTTCAGAGAACTTTTTCATGCTTATGTTATCATGTT 717
 QY 715 AAA--AAGAAGTTACAGTACATCAATTAAGTGCAGTTTAA 754
 DB 718 AACACAGATGTTTACAGTTACCAATTAAGTGCAGTTTAA 760
 RESULT 20
 B0648608 983 bp mRNA linear EST 15-JUL-2002
 LOCUS B0648608
 DEFINITION AGENCOURT_8298360 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6269639
 5', mRNA sequence.
 ACCESSION B0648608
 VERSION B0648608.1 GI:21772780
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 983)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 CDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:

http://image.lnl.gov
 Plate: LICM2443 row: h column: 24
 High quality sequence stop: 563.
 Location/Qualifiers
 1..983
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6269639"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."
 BASE COUNT 227 a 247 c 282 g 227 t
 ORIGIN
 Query Match 81.3%; Score 612.8; DB 13; Length 983;
 Best local similarity 94.3%; Pred. No. 1.2e-161;
 Matches 713; Conservative 0; Mismatches 32; Indels 11; Gaps 7;
 QY 1 ATGGCGAGCGTGGAAAGACGAGAGGAAACCTGCGCTTGCTTCACTCCGGAGCTCA 60
 DB 16 ATGGCGAGCGTGGAAAGACGAGAGGAAACCTGCGCTTGCTTCACTCCGGAGCTCA 75
 QY 61 GGCTCCAACTCGGGAGCGCAACAAGATGTTCTCCCTCAAGAAGTGAACCGGTGGCCATG 120
 DB 76 GGCTCCAACTCGGGAGCGCAACAAGATGTTCTCCCTCAAGAAGTGAACCGGTGGCCATG 135
 QY 121 TGAAGCTGGGAGCGTGAAGTGCATACGTGCGCATCTGAGAGGTCAGGTGATGATGTC 180
 DB 136 TGAAGCTGGGAGCGTGAAGTGCATACGTGCGCATCTGAGAGGTCAGGTGATGATGTC 195
 QY 181 TGTCTTAGATGTCAGCTGAAAAAACAAGAGACGTGTGTGTGCTGGGAGAGATGT 240
 DB 196 TGTCTTAGATGTCAGCTGAAAAAACAAGAGACGTGTGTGTGCTGGGAGAGATGT 255
 QY 241 AATCATTTCTTCCACAACCTGTCATGTCCCTGTGGGTAAACAGAACTATGCTCCCT 300
 DB 256 AATCATTTCTTCCACAACCTGTCATGTCCCTGTGGGTAAACAGAACTATGCTCCCT 315
 QY 301 CTCTGCACAGAGACTGGGTGTGCCAAGAAATCGGCAATGAGAGTGTAGAGGCTTC 360
 DB 316 CTCTGCACAGAGACTGGGTGTGCCAAGAAATCGGCAATGAGAGTGTAGAGGCTTC 375
 QY 361 TTAGGCGAGTGTTCAGAGCCCTGTGTGATCTTGTATTCAGTGCCTTCAAGGCTTAA 420
 DB 376 TTAGGCGAGTGTTCAGAGCCCTGTGTGATCTTGTATTCAGTGCCTTCAAGGCTTAA 435
 QY 421 ACACCTACAGGAGTGAATCTTCAATAGAGACCCGATGATCTGTGTC--TTTGAAGTC 478
 DB 436 ACACCTACAGGAGTGAATCTTCAATAGAGACCCGATGATCTGTGTC--TTTGAAGTC 495
 QY 479 ATCAAGGCTTGCTT--AGCATTTGTCAGTTTATCTTGAAGAAATCTCTGATTTAAGA 536
 DB 496 ATCAAGGCTTGCTT--AGCATTTGTCAGTTTATCTTGAAGAAATCTCTGATTTAAGA 555
 QY 537 AGATATATTTTAAAGTGTCTTCTCACTCTGTGTGTGTGTGCGACACAGCTTA 596
 DB 556 AGATATATTTTAAAGTGTCTTCTCACTCTGTGTGTGTGTGCGACACAGCTTA 615
 QY 597 GAAAGCTATATAAAAAAGGAAAG--CTCCAAATGAATCAAC--TTATATTTAACCATT 654
 DB 616 GAAAGCTATATAAAAAAGGAAAG--CTCCAAATGAATCAAC--TTATATTTAACCATT 675
 QY 655 CTATACACA--GGCAGTGAAGCAGTTTC--GAGACTTTTTCAGATGCTTAT--GGTGTAT 709


```

Db      676 CTATCAACAGGCGAGTGGAGCAGTCCCGAGAACTGTGTCATGCTTATGAGGCTAT 735
Qy      710 CAGTTAAAAAGAAATGTTACAGTAAACAATAAGTG 745
Db      736 CAGCTGAAAAAACATTGCTGCGCCCAATAAAGG 771

RESULT 21
Bg753323
LOCUS   60231740p1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4875489 5',
DEFINITION
MRNA sequence.
ACCESSION
Bg753323
VERSION
Bg753323.1 GI:14063976
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 856)
REFERENCE
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM1756 row: 0 column: 10
High quality sequence stop: 767.
Location/Qualifiers
1..856
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4875489"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/note="Organ: eye; Vector: pOT87; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT
231 a 214 c 211 g 200 t

Query Match 80.8%; Score 609; DB 10; Length 856;
Best Local Similarity 94.1%; Pred. No. 1.3e-160;
Matches 712; Conservative 0; Mismatches 30; Indels 15; Gaps 7;

Qy      1 ATGGCCGACGTGGAAGACGAGAGGAAACCTGCGCTCTCACTCCGGAGCTCA 60
Db      18 ATGGCCGACGTGGAAGACGAGAGGAAACCTGCGCTCTCACTCCGGAGCTCA 77
Qy      61 GGCTCCAACTGGGAGGCGACAAGATGTTCTCCCTCAAGAGTGAACGCGGTGCCATG 120
Db      78 GGCTCCAACTGGGAGGCGACAAGATGTTCTCCCTCAAGAGTGAACGCGGTGCCATG 137
Qy      121 TGGACCTGGAGAGTGAAGTGCATAGCTGGCCATCTGAGAGGTCCAGGTGATGATGC 180
Db      138 TGGACCTGGAGAGTGAAGTGCATAGCTGGCCATCTGAGAGGTCCAGGTGATGATGC 197
Qy      181 TGTCTTAGATGTCAGAGTGAACAAACAAGAGAGTGTGTGTGCTGGGGAATGT 240

```

```

Db      198 TGTCTTAGATGTCAGAGTGAACAAACAAGAGAGTGTGTGTGCTGGGGAATGT 257
Qy      241 AATCATTCCTTCCACAACTGCTGCATGTCCCTGTGAGTGAACAGAAATGCTGCCCT 300
Db      258 AATCATTCCTTCCACAACTGCTGCATGTCCCTGTGAGTGAACAGAAATGCTGCCCT 317
Qy      301 CTCTGCCAGACGACGTGGGTGTCTCCAAAGAAATCGGCAAAAGAGTGTGAAGGCTTC 360
Db      318 CTCTGCCAGACGACGTGGGTGTCTCCAAAGAAATCGGCAAAAGAGTGTGAAGGCTTC 377
Qy      361 TTACGCGAGTGTTCACAGGCCCTGTGATCTTGAATCCAGTCCCTACAAAGCTAGA 420
Db      378 TTACGCGAGTGTTCACAGGCCCTGTGATCTTGAATCCAGTCCCTACAAAGCTAGA 437
Qy      421 ACACACAGGAGGATGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGAATC 478
Db      438 ACACACAGGAGGATGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGAATC 497
Qy      479 ATCAAAGCTTGGTT--AGCATTTGTCAGTTTATCTTCAGAAATTCCTGTGATTA 536
Db      498 ATCAAAGCTTGGTTAGCATTTTGTGATTTATCTTCAGAAATTCCTGTGATTA 557
Qy      537 AGATTAATTATTA--AGGTGCTCTTCTCACTCTGTGTGTGTGTGCGACACACCTT 595
Db      558 AGATTAATTATTA--AGGTGCTCTTCTCACTCTGTGTGTGTGTGCGACACACCTT 617
Qy      596 AGAAGTC-CTATAAAGAGAGAGAGAGCTCCAAATTAATGATC--TTATTAATTA 653
Db      618 AGAAGTCCTATAAAGAGAGAGAGCTCCAAATTAATGATC--TTATTAATTA 677
Qy      654 TCTATACACGAGGAGGAGGAGAGAGCTTTCGACCTTTTGTGATGTC-----TT 707
Db      678 TCTATACACGAGGAGGAGGAGAGAGCTTTCGACCTTTTGTGATGTC-----TT 737
Qy      708 ATCAGTTAA--AAAAGATGTTACAGTACAAATTA 742
Db      738 ATCAGTTAACCAAGAAATGTTACAGTACCAACATA 774

RESULT 22
Bi601470 767 bp mRNA linear EST 07-SEP-2001
LOCUS   603249053p1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5300845 5',
DEFINITION
MRNA sequence.
ACCESSION
Bi601470
VERSION
Bi601470.1 GI:15494409
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 767)
REFERENCE
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM11761 row: j column: 14
High quality sequence stop: 701.
Location/Qualifiers
1..767
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5300845"

```

```

/issue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/notes="Organ: brain; Vector: pBlueScript (modified
pBlueScript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTT-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT      192 a      173 c      214 g      188 t
ORIGIN

```

```

Query Match      80.7%; Score 608.8; DB 12; Length 767;
Best Local Similarity 94.2%; Pred. No. 1.5e-160;
Matches 688; Conservative 0; Mismatches 32; Indels 10; Gaps 5;

1 ATGGCCGAGCTGGAGACGAGAGAACTGGCCCTGACCTCCGGAGCTCA 60
35 ATGGCCGAGCTGGAGACGAGAGAACTGGCCCTGACCTCCGGAGCTCA 94
61 GGCTCCAACTCGGAGAGCGACAAGATGTTCTCCCTCAAGAAGTGAACCGGTGGCCATG 120
95 GGCTCCAACTCGGAGAGCGACAAGATGTTCTCCCTCAAGAAGTGAACCGGTGGCCATG 154
121 TGGAGCTGGAGAGCTGGAGTGCATACGTGGCCATCTGCAAGGTCAGGTGATGATGCC 180
155 TGGAGCTGGAGAGCTGGAGTGCATACGTGGCCATCTGCAAGGTCAGGTGATGATGCC 214
181 TGTCTTAGATGTCAGCTGAGAAACAAGAAGAGCTGTTGTGTGGTGGGAGAAATG 240
215 TGTCTTAGATGTCAGCTGAGAAACAAGAAGAGCTGTTGTGTGGTGGGAGAAATG 274
241 AATCATCTCTTCCACAACACTGCTGCATGTCCTGTGGGTGAACAGAAACATCGTCCCT 300
275 AATCATCTCTTCCACAACACTGCTGCATGTCCTGTGGGTGAACAGAAACATCGTCCCT 334
301 CTCTGCGACGAGAGCTGGGTGCTCCAAAGATGCGCAATGAGAGTGTAGAGGCTTC 360
335 CTCTGCGACGAGAGCTGGGTGCTCCAAAGATGCGCAATGAGAGTGTAGAGGCTTC 394
361 TTAGGCGAGTGTGTCAGAGCCCTGCGATCTGTATCCAGGCCCTCAAGGCTAGA 420
395 TTAGGCGAGTGTGTCAGAGCCCTGCGATCTGTATCCAGGCCCTCAAGGCTAGA 454
421 ACACTACAGGAGATGATTTCTTCAATAGAGCCGATGATCTGTGTC--TTTGGATC 478
455 ACACTACAGGAGATGATTTCTTCAATAGAGCCGATGATCTGTGTC--TTTGGATC 514
479 ATCAAAAGCTTGGTT--AGCATTTGTCAAGTTTATCTTCAAGAAATCTCTGTGATTAGA 536
515 ATCAAAAGCTTGGTT--AGCATTTGTCAAGTTTATCTTCAAGAAATCTCTGTGATTAGA 574
537 AGATATTATTATTA--GTTGTCCTTCTTCACTCTGTGTCGTCGCGGACACAGCTT 595
575 AGATATTATTATTA--GTTGTCCTTCTTCACTCTGTGTCGTCGCGGACACAGCTT 634
596 AGAAGTCTATTAATAAAGAGAGCTCAAAATTTGATCACTTATTAATTTCCATTTTC 655
635 AGAAGTCTATTAATAAAGAGAGAGCTCAAAATTTGATCACTTATTAATTTCCATTTTC 694
656 TA--TACAACAGGAGTGAAGAGCTTGA--GACTTTTTCGATGCTTATGCTTATG 710
695 TATATCAACAGGAGTGAAGAGAGCTTGAAGAGCTTATGCTTATGCTTATGCTTATG 754
711 AGTTAAAAA 720
755 AGCTTTAAAA 764

```

RESULT 23

```

A138342/c
LOCUS      A138342      702 bp      mRNA      linear      EST 13-FEB-1999
DEFINITION G96G03.x1 Soares total fetus Nb2HF8.9w Homo sapiens cDNA clone
IMAGE:1939252.3' similar to WF:R10A10.2 CE12670 ZINC FINGER PROTEIN
; mRNA sequence.
ACCESSION  A138342
VERSION    A138342.1 GI:4075269
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE       Unpublished
JOURNAL     Contact: Robert Strausberg, Ph.D.
COMMENT     Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.jnl.gov) for further information.
            Insert Length: 854 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 455.
            Location/Qualifiers
                1..702
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:1939252"
                /dev_stage="8-9 weeks"
                /lab_host="DH10B"
                /clone_lib="Soares total fetus Nb2HF8.9w"
                /note="Vector: pT73D-Pac (Pharmacia) with a modified
                polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
                was prepared from mRNA obtained from pooled 8-9 week
                (total) fetus material with a Not I - oligo(dT) primer [5'
                TGTTCACATCGAAGTGGAGCGGCGCTTATTTTATTTT 3'].
                Double-stranded cDNA was ligated to Eco RI adaptors
                (Pharmacia), digested with Not I and cloned into the Not I
                and Eco RI sites of the modified pT73 vector. Library
                went through one round of normalization, and was
                constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      189 a      175 c      137 g      201 t
ORIGIN

```

```

Query Match      80.6%; Score 607.4; DB 9; Length 702;
Best Local Similarity 96.0%; Pred. No. 3.5e-160;
Matches 667; Conservative 0; Mismatches 21; Indels 7; Gaps 4;

67 AAGTGGAGAGCGACAAATGTTCTTCCCTCAAGAGTGAACGCGGTGCGATGTGAGC 126
702 AAGTGGAGAGCGACAAATGTTCTTCCCTCAAGAGTGAACGCGGTGCGATGTGAGC 643
127 TGGAGAGTGGAGTGGATACGTCGCGCATCTGCAAGGTCCAGGTATGGATGCTGCTT 186
642 TGGAGAGTGGAGTGGATACGTCGCGCATCTGCAAGGTCCAGGTATGGATGCTGCTT 583
187 AGATGTCAAGCTGAAACAAACAAGAGACTGTGTTGTGTGTCGCGAGATGATCAT 246
582 AGATGTCAAGCTGAAACAAACAAGAGACTGTGTTGTGTGTCGCGAGATGATCAT 523
247 TCTTCCACAATGTGTCATGTCCTGTGGGTGAACAAGAAATCGCTGCTCTGTC 306
522 TCTTCCACAATGTGTCATGTCCTGTGGGTGAACAAGAAATCGCTGCTCTGTC 463
307 CAGCAGAGCTGGGTGTCCTCAAGAAATCGGCAATAGAGTGTGTAAGAGCTTCTTAGG 366
462 CAGCAGAGCTGGGTGTCCTCAAGAAATCGGCAATAGAGTGTGTAAGAGCTTCTTAGG 403
367 CAGTGTTCAGAGCCCTGTGATCTTGTATCCAGTGCCTTCAAGGCTTGAACACTTA 426
402 CAGTGTTCAGAGCCCTGTGATCTTGTATCCAGTGCCTTCAAGGCTTGAACACTTA 343

```

Query Match	80.58;	Score 606.6;	DB 12;	Length 847;
Best Local Similarity	94.28;	Pred. No. 6.3e-160;		
Matches 685;	Conservative 0;	Mismatches 34;	Indels 8;	Gaps 5;
QY	1	ATGGCCGACGAGGAAAGACGAGAGGAAACCTGCGCCTTGCCCTTCACTCCGGGAGCTCA	60	
Db	122	ATGGCCGACGAGGAAAGACGAGAGGAAACCTGCGCCTTGCCCTTCACTCCGGGAGCTCA	181	
QY	61	GGCTCCAACTGGGAGGCGGACAAAGATGTTCTCCCTCAGAAAGTGAACCGGATGGCCATG	120	
Db	182	GGCTCCAACTGGGAGGCGGACAAAGATGTTCTCCCTCAGAAAGTGAACCGGATGGCCATG	241	
QY	121	TGGAGCTGGGACGTGGATGCGGATACGTGGCCATCTGCAAGGCTCCAGGATGATGATGCC	180	
Db	242	TGGAGCTGGGACGTGGATGCGGATACGTGGCCATCTGCAAGGCTCCAGGATGATGATGCC	301	
QY	181	TGCTTTAGATGTCACAGCTGTGAAAACAAACAAGAGACTGTGTGTGCTTGGGAGAAATGT	240	
Db	302	TGCTTTAGATGTCACAGCTGTGAAAACAAACAAGAGACTGTGTGTGCTTGGGAGAAATGT	361	
QY	241	AATCATTCCTTCCAAACTGCTGCATGTTCCTCTGTGGGTGAACAAGAACATTCGTCGCCCT	300	
Db	362	AATCATTCCTTCCAAACTGCTGCATGTTCCTCTGTGGGTGAACAAGAACATTCGTCGCCCT	421	
QY	301	CTTCCGACGAGGACTGGGTGCTCCAAAGAAATCGGCAATGAGATGGTGTGAAGGCTTC	360	

[illegible]

DB	839	Accession	043
RESULT 25			
BG576768			
LOCUS	748 bp	mRNA	linear
DEFINITION	602599091.F1 NIH-MGC_87 Homo sapiens cDNA clone IMAGE:4707838 5'		
ACCESSION			
VERSION	BG576768		
KEYWORDS	BG576768.1 GI:13584421		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 748)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	Mammalian Gene Collection (MGC)		

JOURNAL

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL0578 row: e column: 23

High quality sequence stop: 748.

FEATURES

source

1. 748

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4707838"

/tissue_type="mammary adenocarcinoma, cell line"

/lab_host="NIH MGC 87"

/note="Organ: breast; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC library."

BASE COUNT 201 a 156 c 187 g 203 t 1 others

ORIGIN

Query Match

80.4%; Score 606; DB 10; Length 748;

Best Local Similarity 96.2%; Pred. No. 8.9e-160;

Matches 686; Conservative 0; Mismatches 16; Indels 11; Gaps 6;

50 CCGGAGCTCAGGCTCCAGTCCGGAGGCGCAAGATGTTCTCCCTCAAGAATGGAAG 109

5 CCGGAGCTCAGGCTCCAGTCCGGAGGCGCAAGATGTTCTCCCTCAAGAATGGAAG 64

110 CCGTGGCCATGTGAGCTGGGAGCTGGAAGTGCATGCGGCGCATCTGAGGGTCCAG 169

65 CCGTGGCCATGTGAGCTGGGAGCTGGAAGTGCATGCGGCGCATCTGAGGGTCCAG 124

170 TGATGATGCTGTCTTAGATGTCAAGCTGAAAAACAAGAGAGCTGTGTGTGTCT 229

125 TGATGATGCTGTCTTAGATGTCAAGCTGAAAAACAAGAGAGCTGTGTGTGTCT 184

230 GGGGAAATGTATCATTTCTTCCACAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 289

185 GGGGAAATGTATCATTTCTTCCACAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 244

290 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 349

245 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 304

350 TAGAAGCTTCTTAGAGGAGTGTTCAGAGCCCTGTGTGTGTGTGTGTGTGTGTGT 409

305 TAGAAGCTTCTTAGAGGAGTGTTCAGAGCCCTGTGTGTGTGTGTGTGTGTGTGT 364

410 CAAAGCTTAGAACAACAGAGGAGTGAATTTCTCAATATGAGAGCCATGATGTGTGTC 469

365 CAAAGCTTAGAACAACAGAGGAGTGAATTTCTCAATATGAGAGCCATGATGTGTGTC 424

470 --TTTGAGCTCATCAAGGCTTGGT--AGCATTTGTCAAGTTTATCTTCAAGAAATTC 525

425 CTTTGGAGCTCATCAAGGCTTGGT--AGCATTTGTCAAGTTTATCTTCAAGAAATTC 484

526 TGTGATTAAGAAATATTTATTAAGTGTGCTTCTTCAATCTGTGTGTGTGTGTGTCG 585

485 TGTGATTAAGAAATATTTATTAAGTGTGCTTCTTCAATCTGTGTGTGTGTGTGTCG 544

586 CACACAGCTTAAGAAATATTTATTAAGTGTGCTTCTTCAATCTGTGTGTGTGTGTCG 643

545 CACACAGCTTAAGAAATATTTATTAAGTGTGCTTCTTCAATCTGTGTGTGTGTGTCG 604

Qy 644 TTACCATTTCTATACACAGGAGTGGAGCAGTTTC--GAGCTTTTGCATGCTTA 701

Db 605 TTACCATTTCTAT--ACAGGAGTGGAGCAGTTTCAGAGACTTTTGTGATCTTA 661

Qy 702 TGGTGTATGATTAAGAAATGTTACAGTAACAATTAAGTGCAGTTTAA 754

Db 662 TGGTGTATGATTAAGAAATGTTACAGTAACAATTAAGTGCAGTTTAA 714

RESULT 26

CB990062

LOCUS

CB990062 782 bp mRNA linear EST 01-MAY-2003

AGENCY: NCBI, 13902487 NIH MGC 147 Homo sapiens cDNA clone

IMAGE:30346117 5', mRNA sequence.

CB990062.1 GI:30284582

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Dr. Stefan Hanson

CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help

and advice from Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDAM385 row: b column: 14

High quality sequence stop: 635.

FEATURES

source

1. 782

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30346117"

/tissue_type="Human Placenta"

/lab_host="DH10B TONK"

/clone_lib="NIH MGC 147"

/note="Organ: Placenta; Vector: pBluescript; Site: 1: all-XhoI; Site: 2: BamHI; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5; This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: This is a NIH MGC library."

BASE COUNT

195 a 186 c 215 g 186 t

ORIGIN

Query Match

80.0%; Score 603.4; DB 14; Length 782;

Best Local Similarity 96.6%; Pred. No. 4.9e-159;

Matches 650; Conservative 0; Mismatches 16; Indels 7; Gaps 3;

Qy 1 ATGGCCGAGCTGGAAGACGAGAGAAACCTGCGCTTCTCACTCCGGAGCTTA 60

Db 84 ATGGCCGAGCTGGAAGACGAGAGAAACCTGCGCTTCTCACTCCGGAGCTTA 143

Qy 61 GGCTCAAGTGGGAGGCGACAAGATGTTCTCCCTCAAGAGTGGAGCGGGTGGCCATG 120

Db 144 GGCTCAAGTGGGAGGCGACAAGATGTTCTCCCTCAAGAGTGGAGCGGGTGGCCATG 203

Qy 121 TGGAGCTGGAGCTGGAAGTGCATACGTGCGGCATCTGCAAGGCTCAGGTATGATGCC 180

Db 204 TGGAGCTGGAGCTGGAAGTGCATACGTGCGGCATCTGCAAGGCTCAGGTATGATGCC 263

```

OY 181 TGTCTTAGATGTCAGAGCTGAAAAAACAAGAGAGAGCTGTGTTGCTGTGGGAGAAAT 240
    |||||
Db 264 TGTCTTAGATGTCAGAGCTGAAAAAACAAGAGAGAGCTGTGTTGCTGTGGGAGAAAT 323
    |||||
OY 241 AATCATTTCTTCCACACAGCTGTCGATGTCCTGTGGTGAACAGAACATCGTGCCT 300
    |||||
Db 324 AATCATTTCTTCCACACAGCTGTCGATGTCCTGTGGTGAACAGAACATCGTGCCT 383
    |||||
OY 301 CTCTGCGACAGAGAGCTGTGGTGTCCAAAGAAATGCGCAATGAGACTGTGTGAAGGCTTC 360
    |||||
Db 384 CTCTGCGACAGAGAGCTGTGGTGTCCAAAGAAATGCGCAATGAGACTGTGTGAAGGCTTC 443
    |||||
OY 361 TTAGCCGAGTTGTTCAGAGCCCTGTGTGATCTTGAATCAGATGCGCTTGAACAGGCTTGA 420
    |||||
Db 444 TTAGCCGAGTTGTTCAGAGCCCTGTGTGATCTTGAATCAGATGCGCTTGAACAGGCTTGA 503
    |||||
OY 421 ACACCTACAGAGGAGTGAATCTTCAATAGAGAGCCGATGATCTGTGTC--TTTGGACTC 478
    |||||
Db 504 ACACCTACAGAGGAGTGAATCTTCAATAGAGAGCCGATGATCTGTGTC--TTTGGACTC 563
    |||||
OY 479 ATCAAAGCCTTGCT--AGCATTTGTCAGTTTATCTTGAAGAAATCTGTGATTTAAGA 536
    |||||
Db 564 ATCAAAGCCTTGCTTGATGAGATTTTGTGAGTTTATCTTGAAGAAATCTGTGATTTAAGA 623
    |||||
OY 537 AGATTAATTTATTAAGGTGTCTCTCTACCTGTGTGTGTGTGTGCGCACACAGCTTA 596
    |||||
Db 624 AGATTAATTTATTAAGGTGTCTCTCTACCTGTGTGTGTGTGTGCGCACACAGCTTA 683
    |||||
OY 597 GAAGTGTCTATTAAGAAAGAGAGCTGCAATTTGAATTCACCTTAT--AATTACCCANT 653
    |||||
Db 684 GAAGTGTCTATTAAGAAAGAGAGCTGCAATTTGAATTCACCTTATTTAATTTACCCANT 743
    |||||
OY 654 TCTATACACAGG 666
    |||||
Db 744 TCTATACACAGG 756
    |||||

```

```

RESULT 27
LOCUS BG742338 904 bp mRNA linear EST 15-MAY-2001
DEFINITION 602631707F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4777129 5',
            mRNA sequence.
ACCESSION BG742338
VERSION BG742338.1 GI:14052991
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 904)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cge@bbs-remail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LLM10630 row: m column: 02
            High quality sequence stop: 836.
            Location/Qualifiers
                1..904
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4777129"
                /lab_host="DH10B (T1 phage-resistant)"
                /clone_1ib="NCI_CGAP_Skn3"
FEATURES
    source

```

```

/note="Organ: skin; Vector: pCMV-Sport6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 215 a 223 c 254 g 211 t 1 others
ORIGIN

```

```

Query Match 79.2%; Score 597.2; DB 10; Length 904;
Best Local Similarity 93.1%; Pred. No. 2,9e-157;
Matches 714; Conservative 0; Mismatches 39; Indels 14; Gaps 8;

```

```

OY 1 ATGCCGAGCTGGAGAGAGAGAGAGAAACCTGCGCTGAGCTCTCACTCCGGAGCTCA 60
    |||||
Db 134 ATGCCGAGCTGGAGAGAGAGAGAGAAACCTGCGCTGAGCTCTCACTCCGGAGCTCA 193
    |||||
OY 61 GGCTCCAGCTGGAGAGAGAGAGAGATGTTCTCTCCAGAACTGAAACCGGTGGCCATG 120
    |||||
Db 194 GGCTCCAGCTGGAGAGAGAGAGAGATGTTCTCTCCAGAACTGAAACCGGTGGCCATG 253
    |||||
OY 121 TGGAGCTGGAGAGCTGGAGAGAGAGATGTTCTCTCCAGAACTGAAACCGGTGGCCATG 180
    |||||
Db 254 TGGAGCTGGAGAGCTGGAGAGAGAGATGTTCTCTCCAGAACTGAAACCGGTGGCCATG 313
    |||||
OY 181 TGTCTTAGATGTCAGAGCTGAAAAAACAAGAGAGAGCTGTGTGTGTGTGGGAGAAAT 240
    |||||
Db 314 TGTCTTAGATGTCAGAGCTGAAAAAACAAGAGAGAGCTGTGTGTGTGTGGGAGAAAT 373
    |||||
OY 241 AATCATTTCTTCCACACAGCTGTCGATGTCCTGTGGTGAACAGAACATCGTGCCT 300
    |||||
Db 374 AATCATTTCTTCCACACAGCTGTCGATGTCCTGTGGTGAACAGAACATCGTGCCT 433
    |||||
OY 301 CTCTGCGACAGAGAGCTGTGGTGTCCAAAGAAATGCGCAATGAGACTGTGTGAAGGCTTC 360
    |||||
Db 434 CTCTGCGACAGAGAGCTGTGGTGTCCAAAGAAATGCGCAATGAGACTGTGTGAAGGCTTC 493
    |||||
OY 361 TTAGCCGAGTTGTTCAGAGCCCTGTGTGATCTTGAATCAGATGCGCTTGAACAGGCTTA 420
    |||||
Db 494 TTAGCCGAGTTGTTCAGAGCCCTGTGTGATCTTGAATCAGATGCGCTTGAACAGGCTTA 553
    |||||
OY 421 ACACCTACAGAGGAGTGAATCTTCAATAGAGAGCCGATGATCTGTGTC--TTTGAATC 478
    |||||
Db 554 ACACCTACAGAGGAGTGAATCTTCAATAGAGAGCCGATGATCTGTGTC--TTTGAATC 613
    |||||
OY 479 ATCAAAGCCTTGCT--AGCATTTGTCAGTTTATCTTGAAGAAATCTC--TGATTTAAG 535
    |||||
Db 614 ATCAAAGCCTTGCTTGATGAGATTTTGTGAGTTTATCTTGAAGAAATCTC--TGATTTAAG 673
    |||||
OY 536 AAGATTAATTTATTAAG--GTGTGCTCTTCACTGCTGTGTGTGTGTGCGCACACAGCT 594
    |||||
Db 674 AAGATTAATTTATTAAG--GTGTGCTCTTCACTGCTGTGTGTGTGTGCGCACACAGCT 733
    |||||
OY 595 TAGAAGTGTCTATTAAGAAAGAGAGAGCTGCAATTTGAAT--CACCTTATTTATTC 650
    |||||
Db 734 TAGAAGTGTCTATTAAGAAAGAGAGAGCTGCAATTTGAATTTATTTATTTATTC 793
    |||||
OY 651 --AATTCTATACACAGAGAGAGAGAGCTTCC--GAGCTTTTGTGATGCTTATGTT 706
    |||||
Db 794 CATTCTATACACAGAGAGAGAGAGAGCTTCC--GAGCTTTTGTGATGCTTATGTT 853
    |||||
OY 707 GATCAGTTATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 753
    |||||
Db 854 GATCAGTTATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
    |||||

```

```

RESULT 28
LOCUS CA449255/c 692 bp mRNA linear EST 08-NOV-2002
DEFINITION CA449255
            UI-H-ED0-4ys-n-12-0-UI-s1 NCI CGAP ED0 Homo sapiens cDNA clone
            UI-H-ED0-4ys-n-12-0-UI-3', mRNA sequence.
ACCESSION CA449255
VERSION CA449255.1 GI:24813675
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

ORGANISM	Human sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 692)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT	Unpublished Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Dr. Jose Mercuende cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@iowa.edu Seq primer: M13 FORWARD POLYA=Yes.
FEATURES	<p>Location/Qualifiers</p> <p>1..692</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="UI-H-ED0-ays-n-12-0-U"</p> <p>/tissue_type="Chondrosarcoma"</p> <p>/dev_stage="Adult"</p> <p>/lab_host="DH10B (Life Technologies)"</p> <p>/clone_1ib="NCI CGAP ED0"</p> <p>/note="Organ: Left Pubic Bone; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP ED0 is a cDNA library containing the following tissue(s): Chondrosarcoma cell line C55. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.</p> <p>TAG_Lib=UI-H-ED0</p> <p>TAG_Tissue=chondrosarcoma</p> <p>TAG_SEQ=CGTCAAGGCT"</p>
BASE COUNT	181 a 169 c 135 g 206 t 1 others
ORIGIN	
Query Match	79.0%; Score 595.8; DB 14; Length 692;
Best Local Similarity	97.0%; Pred. No. 6.5e-157;
Matches 650; Conservative	0; Mismatches 13; Indels 7; Gaps 4;
QY	92 CCCTCAAGAAGTGGAAAGCGGGTGCACATGTGGAGCTGGGAGCTGGAGTGCAGTACGCGG 151
Db	692 CCCTCAAGAAGTGGAAAGCGGGTGCACATGTGGAGCTGGGAGCTGGAGTGCAGTACGCGG 633
QY	152 CCATCTGCAGGGTCCAGGTGATGATGATGCTCTTACATGTCAGCTGAAACAAACAAG 211
Db	632 CCACTGAGGGTCCAGGTGATGATGATGCTCTTACATGTCAGCTGAAACAAACAAG 573
QY	212 AGGACTGTGTGTGGTCTGGGGAGAAATGATCAATTCCTTCACAACTGCTGATGTC 271
Db	572 AGGACTGTGTGTGGTCTGGGGAGAAATGATCAATTCCTTCACAACTGCTGATGTC 513
QY	272 TGTGGGTGAACAGAAACAATGCGTGCCTCTCCACAGACAGGACTGGGTGGTCCAAAGAA 331
Db	512 TGTGGGTGAACAGAAACAATGCGTGCCTCTCTCCACAGACAGGACTGGGTGGTCCAAAGAA 453
QY	332 TCGGCAATGAGAGTGTGTAGAAAGCTCTTAGCGCAGATTGTCAGAGCCCTGGTGGATC 391
Db	452 TCGGCAATGAGAGTGTGTAGAAAGCTCTTAGAGGCAATTGTCAGAGCCCTGGTGGATC 393
QY	392 TTGTATATTCAGTGCCTCAAGAGCTGAGACACTACAGGGGATGATTTCTTCAATATGGA 451

[illegible]

ORIGIN	BASE COUNT	178 a	166 c	198 g	174 t
Query Match	78.8%;				
Best Local Similarity	96.1%;				
Matches	664; Conservative	0;	Mismatches	18;	Indels 9; Gaps 5;
Qy	1	ATGGCGAGCTGTGAAAGACGAGAGGAAACCTGCGCCCTGACCTCCAGGAGGACTCA	60		
Db	18	ATGGCGGAGCTGGAAGACGGAAGGAAACCTCGCCCTGACCTCCAGGAGGACTCA	77		
Qy	61	GGCTCCAACTCGGAGGCGACAAAGATGTTCTCTTCAAGAGTGAACGCGGTGGCCATG	120		
Db	78	GGCTCCAAAGTCGGAGAGGCGACAAGATGTTCTCTTCAAGAGTGAACGCGGTGGCCATG	137		
Qy	121	TGGAGCTGGGAGAGTGGAGTGCATACGTGCGGCATCTGAGGGTCCAGTATGATGCTC	180		
Db	138	TGGAGCTGGGAGAGTGGAGTGCATACGTGCGGCATCTGAGGGTCCAGTATGATGCTC	197		
Qy	181	TGCTTAAATGTCAAGCTGAAAAACAACAAGAGACTGTGTGTGATCTGGGAGAAATGT	240		
Db	198	TGCTTAAATGTCAAGCTGAAAAACAACAAGAGACTGTGTGTGATCTGGGAGAAATGT	257		
Qy	241	AATCATCTCTTCCACACACTGCTGCATGTCTCTGTGGGTGAACAGAAACAATGCTGCTCT	300		
Db	258	AATCATCTCTTCCACACACTGCTGCATGTCTCTGTGGGTGAACAGAAACAATGCTGCTCT	317		
Qy	301	CTCTGCCACAGAGACTGGGTGTGTCCAAAGAAATCGGCAATGAGATGGTTGAAGAGCTTC	360		
Db	318	CTCTGCCACAGAGACTGGGTGTGTCCAAAGAAATCGGCAATGAGATGGTTGAAGAGCTTC	377		
Qy	361	TTAGGCGCAGTTGTTCAAGAGCCCTGTGTGATCTTGTATCCAGTGCCTTACAAAGGCTAGA	420		
Db	378	TTAGGCGCAGTTGTTCAAGAGCCCTGTGTGATCTTGTATCCAGTGCCTTACAAAGGCTAGA	437		
Qy	421	ACACTACAGGGGATGAATCTTCAATATGGAGCGCATGGATCTGTGGTC--TTTGGATCTC	478		
Db	438	ACACTACAGGGGATGAATCTTCAATATGGAGCGCATGGATCTGTGGTC--TTTGGATCTC	497		
Qy	479	ATCAAAAGCCTTGTT--AGCATTTGTGCATTTTATCTTCAGAAATCTCTGTGATTAAGA	536		
Db	498	ATCAAAAGCCTTGTTAGATTTTGTGATTTTATCTTCAGAAATCTCTGTGATTAAGA	557		
Qy	537	AGAT--AATTTATTTAAAGTGGTCTCTTCTTACCTCTGTGGTGTGTGTGGCCACACAGCT	594		
Db	558	AGATTAATTTATTTAAAGTGGTCTCTTCTTACCTCTGTGGTGTGTGTGGCCACACAGCT	617		
Qy	595	TAGAAGTGTCTATTAATAA--AAGGAAAGAGCTCCAAATTTGAATCACCTT--ATAATTAACCA	651		
Db	618	TAGAAGTGTCTATTAATAA--AAGGAAAGAGAGCTCCAAATTTGAATCACCTTATTAATTAACCA	677		
Qy	652	TTTCTATATCAACAGGCACTGGAGACAGTTTC	682		
Db	678	TTTCTATATCAACAGGCACTGGAGACAGTTTC	708		
RESULT 30					
LOCUS	BG037022	840 bp	mRNA	linear	EST 24 -JAN-2001
DEFINITION	602287341P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4374443 5',				
VERSION	BG037022		mRNA sequence.		
ACCESSION	BG037022.1	GI:12432833			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 840)				
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished				
	Contact: Robert Strausberg, Ph.D.				

Email: ccgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LRAM0038 row: b column: 12
High quality sequence stop: 806.

location/Qualifiers
1. .840

FEATURES
SOURCE

BASE COUNT	Institutes of Health. Note: this is a NIH_MGC Library."			
ORIGIN	219 a	191 c	232 g	198 t
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1
31	1	1	1	1
32	1	1	1	1
33	1	1	1	1
34	1	1	1	1
35	1	1	1	1
36	1	1	1	1
37	1	1	1	1
38	1	1	1	1
39	1	1	1	1
40	1	1	1	1
41	1	1	1	1
42	1	1	1	1
43	1	1	1	1
44	1	1	1	1
45	1	1	1	1
46	1	1	1	1
47	1	1	1	1
48	1	1	1	1
49	1	1	1	1
50	1	1	1	1
51	1	1	1	1
52	1	1	1	1
53	1	1	1	1
54	1	1	1	1
55	1	1	1	1
56	1	1	1	1
57	1	1	1	1
58	1	1	1	1
59	1	1	1	1
60	1	1	1	1
61	1	1	1	1
62	1	1	1	1
63	1	1	1	1
64	1	1	1	1
65	1	1	1	1
66	1	1	1	1
67	1	1	1	1
68	1	1	1	1
69	1	1	1	1
70	1	1	1	1
71	1	1	1	1
72	1	1	1	1
73	1	1	1	1
74	1	1	1	1
75	1	1	1	1
76	1	1	1	1
77	1	1	1	1
78	1	1	1	1
79	1	1	1	1
80	1	1	1	1
81	1	1	1	1
82	1	1	1	1
83	1	1	1	1
84	1	1	1	1
85	1	1	1	1
86	1	1	1	1
87	1	1	1	1
88	1	1	1	1
89	1	1	1	1
90	1	1	1	1
91	1	1	1	1
92	1	1	1	1
93	1	1	1	1

	Query Match	78.6%	Score 592.8	DB 10	Length 840
	Best Local Similarity	92.7%	Pred. No. 5e-156		
	Matches 701	Conservative 0	Mismatches 42	Indels 13	Gaps 7
QY	1	ATGGCCGACGCTGGAGAAGGAGAAACCTGCGCCCTGCGCTCTCACTCGGAGAGCTCA	60		
Db	29	ATGGCCGACGCTGGAGAAGGAGAAACCTGCGCCCTGCGCTCTCACTCGGAGAGCTCA	88		
QY	61	GGCTCCAAAGTCGGGAGGCGCAAGATGTTCTCCCTCAAGAACTGAAACGGGTGGCAATG	120		
Db	89	GGCTCCAAAGTCGGGAGGCGCAAGATGTTCTCCCTCAAGAACTGAAACGGGTGGCAATG	148		
QY	121	TGGAGCTGGGACGCTGGAGTGCATACGTGCGCCATCTGCAGGGGTCCAGGTGATGATGCC	180		
Db	149	TGGAGCTGGGACGCTGGAGTGCATACGTGCGCCATCTGCAGGGGTCCAGGTGATGATGCC	208		
QY	181	TGCTTAGATGTCAGGCTGAAAACAACAGAGGACTGTGTGGTCTGGGAGAAATGT	240		
Db	209	TGCTTAGATGTCAGGCTGAAAACAACAGAGGACTGTGTGGTCTGGGAGAAATGT	268		
QY	241	AATCATTTCTTCCACAACTGCTGCATGTCCCTGTGGTGAACAGAACTCGCTGCCCT	300		
Db	269	AATCATTTCTTCCACAACTGCTGCATGTCCCTGTGGTGAACAGAACTCGCTGCCCT	328		
QY	301	CTCTGCACGAGACCTGGGTGCTCAAAAGATGGGCAAAATGAGAGTGTAGAAAGGCTTC	360		
Db	329	CTCTGCACGAGACCTGGGTGCTCAAAAGATGGGCAAAATGAGAGTGTAGAAAGGCTTC	388		
QY	361	TTTAGCGAGTTGTTCAAGGCCCTGTGGATCTTGTATCCAGTGCCCTCAAAAAGCTAGA	420		
Db	389	TTTAGCGAGTTGTTCAAGGCCCTGTGGATCTTGTATCCAGTGCCCTCAAAAAGCTAGA	448		
QY	421	ACACTACAGGGGATGAAATTTTCAAAATAGAGCCGATGATCTGTGTGTC--TTTGCATC	478		
Db	449	ACACTACAGGGGATGAAATTTTCAAAATAGAGCCGATGATCTGTGTGTC--TTTGCATC	508		
QY	479	ATCAAAAGCCTTGG--TTAGCATTTGTCAAGTTTATCTTCAGAAATTCCTCTGATTTAGAA	537		
Db	509	ATCAAAAGCCTTGGTTTAGCATTTGTCAAGTTTATCTTCAGAAATTCCTCTGATTTAGAA	568		
QY	538	GATAAATTTA--TTAAAGGTGTCCTTCACTCTGTGGTGTGTGTGCGGCACACAG--TT	595		

Db 569 GATAATTATTTAAAGTGSTGCTCTTCCTACCTCTGTGGTGTGTGGCGACACAGCTTT 628

QY 566 AGAAGTCTATATAAAAAAGAAAAGAG--TCCAATTGAATCAGCTTATA----ATTACC 649

Db 629 AGAAGTGCTATATAAAAAAGAAAAGAGCGCTCCAAATTGGAAATCACCTTTATTAATTTAACCC 688

QY 650 CATTTCTATACACAGGCGAGTGGAGACAGTTTC--GAGACTTTTGCAGTCTTATGTTG 707

Db 689 ATTTCCTATACAAACAGGCGAGTGGAGCAGGTTTCAGAGAGAACTTTTGGCTCTGATGGTTG 748

QY 708 ATCAGTTAAAAAAGAATGTTACAGTAAACAATAAAG 743

Db 749 ATCCAGTTAACAAAGAGAGTGTCCGTACACAAATTAAG 784

RESULT	31
B1828930	
LOCUS	
DEFINITION	B1828930 827 bp mRNA linear EST 04-OCT-2001 60307509221 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5166790 5' , mRNA sequence.
ACCESSION	B1828930
VERSION	B1828930.1 GI:15940480
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 827)	NIH-MGC	http://mgc.nhlbi.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished
Contact:	Robert Strausberg,	Ph.D.		

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
Plate: LRAM1414 row: h column: 23
High quality sequence start: 2
High quality sequence stop: 679.
Location/Qualifiers
1..827

FEATURES		source	Location/Qualifiers
	1..827		/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone_image="5166790" /clone_type="medulla" /lab_host="DH10B" /clone_lib="NIH MGC 119" /note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI; Site 2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (InvaGen). Research Genetics tracking code 013. Note this is a NIH MGC Library."
BASE COUNT	232 a	183 c	227 g
ORIGIN		184 t	1 others

Query Match	78.2%	Score 589.6	DB 12	Length 827
Best Local Similarity	97.5%	Pred. No. 48-155		
Matches 653; Conservative	0	Mismatches 9	Indels 8	Gaps 5

Qy	1	ATGCGCCGACGTCGAAAGACGAGAGGAAACCTCGCCCTTCCTCACTCCGAGAGCTCA	60
Db	13	ATGGCGGACGTGGAAGACGAGAGGAAACCTTCGCGCCCTTCCTCACTCCGAGAGCTCA	72

QY	61	GGGCTCAAGTGGGAGGGGACAAGATGTTTCTCCCTCAGAAAGTGGAAACGGGTGGCCAG	120
Db	73	GGCTCAAGTGGGAGGGGACAAGATGTTTCTCCCTCAGAAAGTGGAAACGGGTGGCCAG	132
QY	121	TGAGAGTGGAGCGTGGAGTGGCATACGTGGCCCATCTGCAGGGTCCAGGTGATGATGGC	180
Db	133	TGAGAGTGGAGCGTGGAGTGGCATACGTGGCCCATCTGCAGGGTCCAGGTGATGATGGC	192
QY	181	TGTCTTAATGTCAAGCTGAAAAACAACAAGGACTGTGTTGTGTGTCTGGGAGAAATG	240
Db	193	TGTCTTAATGTCAAGCTGAAAAACAACAAGGACTGTGTTGTGTGTCTGGGAGAAATG	252
QY	241	AATCATTCCTTCCACAACCTGCTGCATGTCCCTGGTGGTGAACAGAACAAATGCGTCCCT	300
Db	253	AATCATTCCTTCCACAACCTGCTGCATGTCCCTGGTGGTGAACAGAACAAATGCGTCCCT	312
QY	301	CTCTGCCAGCAGGACTGGGTGGTCCAAAGAAATCGGCAAAATGAGAGTGGTTAGAGGCTTC	360
Db	313	CTCTGCCAGCAGGACTGGGTGGTCCAAAGAAATCGGCAAAATGAGAGTGGTTAGAGGCTTC	372
QY	361	TTAGGGCAGTGTTCAGAGCCCTGGTGGATCTGTGAATTCAGATGGCCCTACAAAGGCTA	420
Db	373	TTAGGGCAGTGTTCAGAGCCCTGGTGGATCTGTGAATTCAGATGGCCCTACAAAGGCTA	432
QY	421	AACATCACGGGAGTAATTTCTTCAAAATAGAGCCGATGATCTGAGTC--TTTGGACTC	478
Db	433	AACATCACGGGAGTAATTTCTTCAAAATAGAGCCGATGATCTGAGTC--TTTGGACTC	492
QY	479	ATCAAAAGCCTTGCTT--AGCATTTGTCACTTTATCTTCAGAAATTTCTGTGATTTAGA	536
Db	493	ATCAAAAGCCTTGCTT--AGCATTTGTCACTTTATCTTCAGAAATTTCTGTGATTTAGA	552
QY	537	AGATTAATTTAAAGGTGATCTTCCACACCTGTGGTGTGTGCGGCACACAGCTTA	596
Db	553	AGATTAATTTAAAGGTGATCTTCCACACCTGTGGTGTGTGCGGCACACAGCTTA	612
QY	597	GAAGTGCTATAAA--AAAGAAAGAGCTTC--AAATTGAATCAC--TTATTAATTTAACCA	652
Db	613	GAAGTGCTATAAA--AAAGAAAGAGCTTC--AAATTGAATCAC--TTATTAATTTAACCA	672
QY	653	TTCTATACAA 662	
Db	673	TTCTATACCA 682	

RESULT	32
LOCUS	B1858784
DEFINITION	B1858784 675 bp. mRNA linear EST 10-OCT-2001
ACCESSION	603386387F1 NIH_MGC_87 Homo sapiens CDNA clone IMAGE:5395388 5 ,
VERSION	mRNA sequence.
KEYWORDS	B1858784
SOURCE	B1858784.1 GI:15999531
ORGANISM	EST.
REFERENCE	Homo sapiens (human)
AUTHORS	Homo sapiens
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL	Mammalia; Euthetia; Primates; Carnathini; Homnidae; Homo.
COMMENT	1 (bases 1 to 675) NIH-MGC http://mhc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: DCTD/DTP CDNA Library Preparation: Life Technologies, Inc. Genomic Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov Plate: LLM12007 row: m column: 21 High quality sequence stop: 675. Location/Qualifiers
FEATURES	

source

1. 675

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5395388"

/issue_type="mammary adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH_MGC_87"

/note="Organ: breast; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

BASE COUNT 174 a 151 c 184 g 166 t

ORIGIN

Query Match 78.1%; Score 588.8; DB 12; Length 675;

Best Local Similarity 97.9%; Pred. No. 6.1e-155;

Matches 652; Conservative 0; Mismatches 2; Indels 12; Gaps 5;

QY 1 ATGGCCGACGTGGAGACGAGAGAGAAACCTGCGCCCTCTCACTCCGGAGCTCA 60

DB 8 ATGGCCGACGTGGAGAGAGAGAGAAACCTGCGCCCTG -CTCACTCCGGAGCTCA 66

QY 61 GGCTCCAGTCGGGAGGCGACAGATGTTCTCCCTCAAGAGTGGAGCGCGTGCCTATG 120

DB 67 GGCTCCAGTCGGGAGGCGACAGATGTTCTCCCTCAAGAGTGGAGCGCGTGCCTATG 126

QY 121 TGGAGCTGGGACGTGGAGTGCATGCTGCGCCATCTGCGAGGCTCAGGTGATGATGCC 180

DB 127 TGGAGCTGGGACGTGGAGTGCATGCTGCGCCATCTGCGAGGCTCA-----GATGCC 179

QY 181 TGCTTTAGATGTCAAGCTGAAAAACAAGAGACTGTGTGTGTGCTGGGAGATGT 240

DB 180 TGCTTTAGATGTCAAGCTGAAAAACAAGAGACTGTGTGTGTGCTGGGAGATGT 239

QY 241 AATCATTTCTTCCACAACCTGCTGCATGTCCTCTGGGTGAAAACAGAACATGCTGCCCT 300

DB 240 AATCATTTCTTCCACAACCTGCTGCATGTCCTCTGGGTGAAAACAGAACATGCTGCCCT 299

QY 301 CTCTCCAGCAGAGCTGGGTGTCCAAAGAAATCGGCAATGAGAGTGTAGAGGCTTC 360

DB 300 CTCTCCAGCAGAGCTGGGTGTCCAAAGAAATCGGCAATGAGAGTGTAGAGGCTTC 359

QY 361 TTAGGCGAGTGTTCAGAGCCCTGTGTGATCTTGAATCCAGTGCCTCAAAAGCTAGA 420

DB 360 TTAGGCGAGTGTTCAGAGCCCTGTGTGATCTTGAATCCAGTGCCTCAAAAGCTAGA 419

QY 421 ACACACAGGGAGTGAATTTCTCAATAGAGCCGATGATCTGTGTGTC--TTTGACTC 478

DB 420 ACACACAGGGAGTGAATTTCTCAATAGAGCCGATGATCTGTGTGTC--TTTGACTC 479

QY 479 ATCAAGAGCTTGG--TTAGCATTTGTCAAGTTTATCTTCAAGAAATCTCTGATTAAGA 537

DB 480 ATCAAGAGCTTGGTTAGCATTTGTCAAGTTTATCTTCAAGAAATCTCTGCAATTAAGA 539

QY 538 GATAATTTATTAAGGTGTCTCTCTCACTGTGTGTGTGTGTGCGGACACAGCTTAG 597

DB 540 GATAATTTATTAAGGTGTCTCTCTCACTGTGTGTGTGTGTGCGGACACAGCTTAG 599

QY 598 AAGTCTATAAAGAGAGAGAGCTCCAAATGATCAGC--TTATTAATTACCATTTCT 656

DB 600 AAGTCTATAAAGAGAGAGAGCTCCAAATGATCAGCCTTTATTAATTACCATTTCT 659

QY 657 ATACAA 662

DB 660 ATACAA 665

RESULT 33

LOCUS BG111792 947 bp mRNA linear EST 30-JAN-2001

DEFINITION 602285379p1 NIH_MGC_86 Homo sapiens CDNA clone IMAGE:4372819 5',

ACCESSION BG111792

VERSION BG111792.1 GI:12605298

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 947)

AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Plate: LLM1003 row: n column: 20

High quality sequence stop: 682.

FEATURES

SOURCE

1. 947

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4372819"

/issue_type="osteosarcoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH_MGC_86"

/note="Organ: bone; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.53 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

BASE COUNT 228 a 241 c 267 g 210 t 1 others

ORIGIN

Query Match 77.9%; Score 587.2; DB 10; Length 947;

Best Local Similarity 97.3%; Pred. No. 2e-154;

Matches 640; Conservative 0; Mismatches 13; Indels 5; Gaps 4;

QY 1 ATGGCCGACGTGGAGACGAGAGAGAAACCTGCGCCCTCTCACTCCGGAGCTCA 60

DB 54 ATGGCCGACGTGGAGAGAGAGAGAAACCTGCGCCCTCTCACTCCGGAGCTCA 113

QY 61 GGCTCCAGTCGGGAGGCGACAGATGTTCTCCCTCAAGAGTGGAGCGCGTGCCTATG 120

DB 114 GGCTCCAGTCGGGAGGCGACAGATGTTCTCCCTCAAGAGTGGAGCGCGTGCCTATG 173

QY 121 TGGAGCTGGGACGTGGAGTGCATGCTGCGCCATCTGCGAGGCTCAGGTGATGATGCC 180

DB 174 TGGAGCTGGGACGTGGAGTGCATGCTGCGCCATCTGCGAGGCTCAGGTGATGATGCC 233

QY 181 TGCTTTAGATGTCAAGCTGAAAAACAAGAGACTGTGTGTGTGCTGGGAGATGT 240

DB 234 TGCTTTAGATGTCAAGCTGAAAAACAAGAGACTGTGTGTGTGCTGGGAGATGT 293

QY 241 AATCATTTCTTCCACAACCTGCTGCATGTCCTCTGGGTGAAAACAGAACATGCTGCCCT 300

DB 294 AATCATTTCTTCCACAACCTGCTGCATGTCCTCTGGGTGAAAACAGAACATGCTGCCCT 353

QY 301 CTCTCCAGCAGAGCTGGGTGTCCAAAGAAATCGGCAATGAGAGTGTGATGAGGCTTC 360

DB 354 CTCTCCAGCAGAGCTGGGTGTCCAAAGAAATCGGCAATGAGAGTGTGATGAGGCTTC 413

QY 361 TTAGCGAGTGTTCAGAGCCCTGTGTGATCTTGAATCCAGTGCCTCAAAAGCTAGA 420

DB 414 TTAGCGAGTGTTCAGAGCCCTGTGTGATCTTGAATCCAGTGCCTCAAAAGCTAGA 473

QY 421 ACACACAGGGAGTGAATTTCTCAATAGAGCCGATGATCTGTGTGTC--TTTGACTC 478

Db	474	ACACTCAGCGGATGAATTCTTCAATAGAGCCCATGTGATCTTGTCCTTTGGAACTC	533
OY	479	ATCAAAGCCTTGG-TTAGCATTGTGCAGTTTATCTTGAAAATTCCTGTGATTAGAA	537
Db	534	ATCAAGCCTTGTTAGCATTTGTGCAGTTTATCTTGAAAATTCCTGGGATTAGAA	593
OY	538	GATAATTTAATAAGCTGTCCTTCTCTACCTCTGTGTGTGTG-TCGGCACACAG-CTT	595
Db	594	GATAATTTAATAAGCTGTCCTTCTCTACCTCTGTGTGTGTGTGTCGGCACACAGCCTT	653
OY	596	AGAAGTGCCTATAAAAAAGAAAGAGCTCCAATTGAATCACCTTAATTATACCATT	653
Db	654	AGAACTGCTATAAAAAGGAGAGAGCTCCAAATTGCATGACTTTATTAATACCAATT	711
RESULT 34			
B0632594			
LOCUS			
DEFINITION	B0632594	629 bp	mRNA linear EST 02-JUL-2002
ACCESSION	i125c04.y1	HR85	islet Homo sapiens cDNA clone IMAGE:6031086 5'
VERSION	B0632594		similar to TR:Q9YSM7 Q9YSM7 RING FINGER PROTEIN. [1] ; mRNA sequence.
KEYWORDS	B0632594.1	GI:21684112	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Unpublished		
JOURNAL	Other ESTs: i125c04.x1		
COMMENT	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Endocrine Pancreas Consortium Harvard University, Howard Hughes Medical Institute Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138 Tel: 617-495-1812 Fax: 617-495-8557 Email: dmelton@biohp.harvard.edu Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu) Seq primer: -40RP from Gibco High quality sequence atbp: 495. Location/Qualifiers 1..629 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6031086" /tissue_type="Purified pancreatic islet" /lab_host="DH10B" /clone_id="HR85 islet" /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI, Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact Information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permut Lab) / Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."		
BASE COUNT	157 a	141 C	173 G
ORIGIN	158 t		

	Query Match	77.2%	Score 582.4	DB 13	Length 629;	
	Best Local Similarity	98.4%	Pred. No.3.8e-153;			
	Matches	610	Conservative	0	Mismatches	6
					Indels	4
					Gaps	2
QY	24	GGAAACCTGCCCCCTGGCCTTCACTCCGGAGGTCAAGGCTCAAAGTCGGAGGCGACA	83			
Db	9	GGAAACCTGGCCCTGGGCTCTCACTCCGGAGGTCAAGGCTCAAAGTCGGAGGCGACA	68			
QY	84	GATGTTCTCCTCAAGAAGTGGAACGGGCTGGCCATGTGAGCTGGGACGTGAGTGC	143			
Db	69	GATGTTCTCCTCAAGAAGTGGAACGGGCTGGCCATGTGAGCTGGGACGTGAGTGC	128			
QY	144	TACGGGCCCATCTGCAAGGGGTCCAAGGTATGGAAGCCGTCTTAGATGTCAAGCTGAAA	203			
Db	129	TACGGGCCCATCTGCAAGGGGTCCAAGGTATGGAAGCCGTCTTAGATGTCAAGCTGAAA	188			
QY	204	CAACCAAGAGGACTGTGTTGTGTCTGGGGAGAAATGTAATCATTTCTTCCCAACTGCTG	263			
Db	189	CAACCAAGAGGACTGTGTTGTGTCTGGGGAGAAATGTAATCATTTCTTCCCAACTGCTG	248			
QY	264	CATGTCCTGTGGGTGAACCAAAATCGTCCCTCTCTGCAAGCAAGACTGGGTGT	323			
Db	249	CATGTCCTGTGGGTGAACCAAAATCGTCCCTCTCTGCAAGCAAGACTGGGTGT	308			
QY	324	CCAAAGAAATGGCAAAATGAGAGGTGTAAGAAGGCTCTTAGCGAGTTGTCAGAGCCCT	383			
Db	309	CCAAAGAAATGGCAAAATGAGAGGTGTAAGAAGGCTCTTAGCGAGTTGTCAGAGCCCT	368			
QY	384	GGTGGATCTTGTAATCCAGTGCCTTACAAAGGCTAGAACACTAGAGGGATGAATTCTTC	443			
Db	369	GGTGGATCTTGTAATCCAGTGCCTTACAAAGGCTAGAACACTAGAGGGATGAATTCTTC	428			
QY	444	AAATBGAAGCCGATGAGTCTGTGGTC--TTTGGACTCATCAAAGCCTTGGTT--AGCAT	499			
Db	429	AAATBGAAGCCGATGAGTCTGTGGTC--TTTGGACTCATCAAAGCCTTGGTT--AGCAT	488			
QY	500	TGTCAAGTTAATCTTCGAATAATCTCTGTGATTAAGAAGATTAATTAAGAAGTGTCC	559			
Db	489	TGTCAAGTTAATCTTCGAATAATCTCTGTGATTAAGAAGATTAATTAAGAAGTGTCC	548			
QY	560	TTCTTACCTCTGTGTGTGTGTGTGCGCACACAGCTTAGAAGTGTCTATAAAAAGAAAGA	619			
Db	549	TTCTTACCTCTGTGTGTGTGTGTGCGCACACAGCTTAGAAGTGTCTATAAAAAGAAAGA	608			
QY	620	GCTCCAAATTGAATCACCCTT	639			
Db	609	GCTCCAAATTGAATCACCCTT	628			
RESULT 35						
LOCUS	Bg121625	745 bp	mRNA	linear	EST 30.-JAN-2001	
DEFINITION	602351581.F1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:4449693 5'					
VERSION	Bg121625					
KEYWORDS	Bg121625.1 GI:12615134					
SOURCE	EST.					
ORGANISM	Homo sapiens (human)					
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/					
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: cgabbs@email.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLW) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be					

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LLM10234 row: a column: 22
High quality sequence stop: 653.
Location/Qualifiers

FEATURES

source

1. 745

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:449693"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: Not1; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."

BASE COUNT

196 a 157 c 209 g 183 t

Query Match

Best Local Similarity 93.5%; Score 582.4; DB 10; Length 745;
Pred. No. 4.1e-153;
Matches 675; Conservative 0; Mismatches 36; Indels 11; Gaps 6;

3 GGGCCAGCTGGAGAGAGAGAGAACTGGCCCTCGCTCTCACTCGGAGCTCAGG 62
19 GCCCAGCTGGAGAGAGAGAGAACTGGCCCTCGCTCTCACTCGGAGCTCAGG 77
63 CTCGAGCTGGAG 122
78 CTCGAGCTGGAG 137
123 GAGCTGGAG 182
138 GAGCTGGAG 197
163 TCTTGAAGTTCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 242
198 TCTTGAAGTTCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 257
243 TCATTCCTTCCACAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 302
258 TCATTCCTTCCACAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 317
303 CTCGAG 362
318 CTCGAG 377
363 AGCGAGCTTCAG 422
378 AGCGAGCTTCAG 437
423 ACTACAGGAG 480
438 ACTACAGGAG 497
481 CAAGCCTTGGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
498 CAAGCCTTGGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 557
541 AATTATTAAG 600
558 AATTATTAAG 616
601 TGTCTATTAAG 660
617 TGTCTATTAAG 671
661 AACAGGAG 720
672 AACAGGAG 729

QY 721 GA 722
Db 730 GA 731

RESULT 36

BE397308 734 bp mRNA linear EST 21-JUL-2000
LOCUS 601288640F1 NIH_MGC_8 Homo sapiens CDNA clone IMAGE:3619424 5',
DEFINITION mRNA sequence.

ACCESSION BE397308
VERSION BE397308.1 GI:3942673
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS 1 (bases 1 to 734)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rudin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LMC289 row: o column: 09
High quality sequence stop: 667.

FEATURES

source

1. 734

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3619424"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_8"
/note="Organ: lymph; Vector: pOT87; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT

203 a 156 c 187 g 188 t

Query Match

Best Local Similarity 93.4%; Score 581.8; DB 10; Length 734;
Pred. No. 6e-153;
Matches 666; Conservative 0; Mismatches 32; Indels 15; Gaps 5;

37 CTCGCTTCTCACTTCGGAGCTCAGGCTCCAGTTCGGAGAGAGAGAGAGAGAG 96
1 CTCGCTTCTCACTTCGGAGCTCAGGCTCCAGTTCGGAGAGAGAGAGAGAGAG 60
97 AAGAAGTGAAG 156
61 AAGAAGTGAAG 120
121 TGCGAGGTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 173
157 TGCGAGGTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 216
121 TGCGAGGTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 173
217 TGT 276
174 TGT 233
277 GTGAAACAGAACTTCCTGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 336

Db 234 GTGAAACAGAACATCGTCGCCCTCTCTGCCACAGACTGGGTGTCMAAGATCGC 293

QY 337 AAATGAGAGGTGTTAGAGGCTTCTAGCCACTTGTTCAGACCCTGGTGATCTTGT 396

Db 294 AAATGAGAGGTGTTAGAGGCTTCTAGCCACTTGTTCAGACCCTGGTGATCTTGT 353

QY 397 ATCCAGTGCCTCAAAAGGCTAGACACTACAGGGAGTGAATCTTCAATAGAGCCGA 456

Db 354 ATCCAGTGCCTCAAAAGGCTAGACACTACAGGGAGTGAATCTTCAATAGAGCCGA 413

QY 457 TGGATCTGTGTC- -TTTGACTCATCAAGCCTTGTT- -AGCATTTGTCACTTTATC 512

Db 414 TGGATCTGTGTCCTTTGGGACTCATCAAGCCTTGTTAGCATTTTGTCACTTTATC 473

QY 513 TTCAGAAATCTCTGTGATTAAGAGATATTTATTAAGGTGTCTTCTCACTCTGT 572

Db 474 TTCAGAAATCTCTGTGATTAAGAGATATTTATTAAGGTGTCTTCTCACTCTGT 533

QY 573 GGTGTGTGTCGCACACAGCTTAGAAGTCTATTAAGAAAGAGAGCTCCAAATTGAA 632

Db 534 GGTGTGTGTCGCACACAGCTTAGAAGTCTATTAAGAAAGAGAGCTCCAAATTGAA 593

QY 633 TCACCT- -TATAATTTACCATTTCTATACACAGGAGTGAAGCACTTTGAG- ACT 688

Db 594 TCACCTTTATAATTTACCATTTCTATACACAGGAGTGAAGCACTTTGAGACT 653

QY 689 TTTTCATGCTTATGTGTGATCAAGTAAAAAGAGTTCACATCAATTA 741

Db 654 TTTTCATGCTTATGTGTGATCAAGTAAAAAGAGTTCACATCAATTA 706

RESULT 37
B1757881 864 bp mRNA linear EST 25-SEP-2001
DEFINITION .603030486F1 NIH_MGC_114 Homo sapiens CDNA clone IMAGE:520676 5',
B1757881
mRNA sequence.
ACCESSION B1757881
VERSION B1757881.1 GI:15749459
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 864)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.liml.gov
plate: LHM1502 row: 1 column: 21
High quality sequence stop: 667.
Location/Qualifiers
1. 864
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:520676"
/lab_host="DH10B"
/clone_11b="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber

BASE COUNT 221 a 207 c 241 g 195 t
ORIGIN
Query Match
Best Local Similarity 92.3%; Pred. No. 4,7e-151;
Matches 707; Conservative 0; Mismatches 43; Indels 16; Gaps 9;
(Invertegen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."
1 ATGCCGACGTGGAAGACGGAAGAAACCTGCGCCCTTGCTCTCACTCCGGAGCTCA 60
Db 76 ATGCCGACGTGGAAGACGGAAGAAACCTGCGCCCTTGCTCTCACTCCGGAGCTCA 135
QY 61 GGCTCCAACTCGGGAGCGCAAGATGTTCTCCCTCAAAAGTGAACCGCGTGGCCAG 120
Db 136 GGCTCCAACTCGGGAGCGCAAGATGTTCTCCCTCAAAAGTGAACCGCGTGGCCAG 195
QY 121 TGGAGCTGGGAGCGTGAAGTGCATACGTGCGCATCTGACAGGATCCAGTGAATGCC 180
Db 196 TGGAGCTGGGAGCGTGAAGTGCATACGTGCGCATCTGACAGGATCCAGTGAATGCC 255
QY 181 TGTCTTAGATGTCAGCTGAAGAAACAAAGAGACTGTGTTGTGCTGGGAGATGT 240
Db 256 TGTCTTAGATGTCAGCTGAAGAAACAAAGAGACTGTGTTGTGCTGGGAGATGT 315
QY 241 AATCATCTCTTCCACAACTGTCGATGTCCTGCTGGGAGAAACAGAAATGCTGCCCT 300
Db 316 AATCATCTCTTCCACAACTGTCGATGTCCTGCTGGGAGAAACAGAAATGCTGCCCT 375
QY 301 CTCTGCACAGAGACTGGTGTGCCAAGAAATGCGCAATGAGAGTGAAGAGCTTC 360
Db 376 CTCTGCACAGAGACTGGTGTGCCAAGAAATGCGCAATGAGAGTGAAGAGCTTC 435
QY 361 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTTGATTCAGTCCCTTCAAAAGCTTCA 420
Db 436 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTTGATTCAGTCCCTTCAAAAGCTTCA 495
QY 421 ACACCTACAGGGAGTGAATCTTCAATAGAGCCGATGGATCTGTGTC- -TTTGGACTC 478
Db 496 ACACCTACAGGGAGTGAATCTTCAATAGAGCCGATGGATCTGTGTC- -TTTGGACTC 555
QY 479 ATCAAGCCTTGTTAGCATTTGTGATGTTTCTTACAGAAA- -TTCTGTGATTAAGAA 537
Db 556 ATCAAGCCTTGTTAGCATTTGTGATGTTTCTTACAGAAA- -TTCTGTGATTAAGAA 615
QY 538 GATTAATTTAT- -AAAGTGTCTTCTCACTCTGTGTGTGTGTCGACACAGCTT- 595
Db 616 GATTAATTTAT- -AAAGTGTCTTCTCACTCTGTGTGTGTGTCGACACAGCTT- 675
QY 596 AAGAGTGTAT- -AAAAAGAGAAAGCTCCAAA- -TTGAATCACCTTA- -AATTAAC 649
Db 676 AAGAGTGTAT- -AAAAAGAGAAAGCTCCAAA- -TTGAATCACCTTA- -AATTAAC 735
QY 650 CATTTCTATACAAAGGAGCT- -GGAAGCAGTTTC- -GAGACTTTTGTGATGCTTATGG 704
Db 736 CATTTCTATACAAAGGAGCT- -GGAAGCAGTTTC- -GAGACTTTTGTGATGCTTATGG 795
QY 705 TGGATCAGTTAAAGAAATGTTAAGTAAATTAAGTGCAGTT 750
Db 796 TGGATCAGTTAAGAAATGTTAAGTAAATTAAGTGCAGTT 841
RESULT 38
Bg716023 793 bp mRNA linear EST 08-MAY-2001
LOCUS Bg716023
DEFINITION .602677415F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:4800170 5',
Bg716023
mRNA sequence.
ACCESSION Bg716023
VERSION Bg716023.1 GI:13995210
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (basses 1. to 793)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.

FEATURES	Location/Qualifiers
source	1. .793

Query Match	76.2%	Score 574.8	DB 10	Length 793
Best Local Similarity	95.0%	Pred. No. 5.9e-151		
Matches 684	Conservative	0	Mismatches 17	Indels 19
				Gaps 8

QY	1	ATGGCCGACGAGGAAACGAGAGGAAACCTGGGCCCTGGACCTCTCACTCCGGAGAGCTCA	60
Db	80	ATGGCCGACGAGGAAACGAGAGGAAACCTGGGCCCTGGACCTCTCACTCCGGAGAGCTCA	139
QY	61	GGCTTCAAGTCGGAGGCGCAACAAGTGTCTTCCCTCAAGAAATGGAACGGCGTGGCCATG	120
Db	140	GGCTTCAAGTCGGAGGCGCGCAACAAGTGTCTTCCCTCAAGAAATGGAACGGCGTGGCCATG	199
QY	121	TGAGAGCTGGGACGTGGAGTCGATCGTGGCCCATCTGCAGAGTCCAGAGTATGATGCC	180
Db	200	TGAGAGCTGGGACGTGGAGTCGATCGTGGCCCATCTGCAGAGTCCAGAGTATGATGCC	252
QY	181	TGCTTAGATGTCAAGCTGAAACAAACAAGAGCATGTGTGTGTGCTGTGGAGAAATGT	240
Db	253	TGCTTAGATGTCAAGCTGAAACAAACAAGAGCATGTGTGTGTGCTGTGGAGAAATGT	312
QY	241	AATATATCTTCCCAACAATGCTGCATGTCTCTGTGGGTGAAACAGAACAAATGCTGCCCT	300
Db	313	AATATATCTTCCCAACAATGCTGCATGTCTCTGTGGGTGAAACAGAACAAATGCTGCCCT	372
QY	301	CTCTGCAGCAGCACTGGTGGTCCAAAGAAATCGGCAATGAGATGTGTTGAAGAGCTTC	360
Db	373	CTCTGCAGCAGCACTGGTGGTCCAAAGAAATCGGCAATGAGATGTGTTGAAGAGCTTC	432
QY	361	TTAGCGCATGTTTCAAGACCTCGTGGATCTTGTATCCAGTGCCTTACAAAGGCTAGA	420
Db	433	TTAGCGCATGTTTCAAGACCTCGTGGATCTTGTATCCAGTGCCTTACAAAGGCTAGA	492
QY	421	ACACTACAGGGGATGAATTTCTTCAATTGGAGCGCATGGAATCTGTGGTCC--TTTGGATTC	478
Db	493	ACACTACAGGGGATGAATTTCTTCAATTGGAGCGCATGGAATCTGTGGTCCCTTTTGGAGCTC	552

QY	479	ATCAAGCGCTGGT - AGCAATTGTCAGTTTAACTTCACGAATAATTCGTGATTAAGA	536
Db	553	ATCAGACCCTGGTTAGCATTTTGACAGTTTTAATCTTCAGAAATTCCTCGATTAAAGA	612
QY	537	A-GATTAATTATTA-AGATGGTCCCTCACTCTGTGTGTGTGTGCACACAGCT	594
Db	613	ACGATAATTATTAACCGGTGGTCTCTCACTCTGTGTGTGTGTGCCACACAGCT	672
QY	555	TGAAAGTGCCTAAAAAGAAA--GAGCTCCAATTGAATCACCT-TTATATTTACCC	650
Db	673	TAGAAGTCTCTAAACAACAGAACACGAGCTCCAATTAATCACTTATATTTACCC	732
QY	651	ATTTCTATACAAACGGCAGTGGAGCAATTC - GAGACTTTTTCAGATCTATGGTGA	708
Db	733	ATTTCTATACAAACGGCAGTGGAGCAATTCCTACAGAACTTTTCATAGCTTATATGGTGA	792

RESULT 39	651 bp	mRNA	linear	EST 08-MAY-2001
BG714665				
LOCUS	BG714665	602677016r1	NIH_MGC_96	Homo sapiens cDNA clone IMAGE:479834 5',
DEFINITION				mRNA sequence.
ACCESSION	BG714665			
VERSION	BG714665			
KEYWORDS	EST.	GI:13993596		
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens			
				(human)

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D.

FEATURES
source

BASE COUNT	157 a	148 c	190 g	156 t	
ORIGIN					
Query Match		75.8%	Score 571.8;	DB: 10;	Length 651;
Best Local Similarity		98.1%	Pred. NO. 3.8e-150;		
Matches 611;		Conservative 0;	Mismatches 7;	Indels 5;	Gaps 3.
Qy	1	ATGGCCGACGCTGGAGAGACGAGAGAAACCTGCGCCCTGAGCTTCTCACTCCGGAGCTCA	60		

Db 41 AAAGAATGTTACAGTAACAATAAAGTCAGTTTAA 4

RESULT 41
BU601136 678 bp mRNA linear EST 20-SEP-2002
LOCUS AGENCOURT 10018944 NIH_MGC 142 Homo sapiens cDNA clone
DEFINITION IMAGE:6495009 5', mRNA sequence.
ACCESSION BU601136
KEYWORDS BU601136.1 GI:23252895
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 678)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM2672 row: c column: 10
High quality sequence stop: 527.
Location/Qualifiers
1..678
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6495009"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH MGC 142"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccatattgccc);
Site 2: SfiI (ggccgcctcgccc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-AAGCAGTGTATCAACGACAGTGCATTTACGCCGG-3' and
5'-ATTCTAGAGCGGAGCGCGGCACATG-dT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the >0.5 kb
size fraction (other fractions present in NIH_MGC 141).
Library created in the laboratory of M. Brownstein (NIH,
NIH). Note: this is a NIH MGC library."

BASE COUNT 185 a 129 c 160 g 178 t 26 others

ORIGIN

Query Match 74.2%; Score 559.4; DB 13; Length 678;
Best local Similarity 96.8%; Pred. No. 1.2e-146;
Matches 613; Conservative 0; Mismatches 13; Indels 7; Gaps 4;

Db 129 GACAGTGAAGTGCATGATGCGGCATCTGCAGGCGCCAGGTAGTGCCTGCTTAG 188
|||||
Db 1 GGAAGTGAAGTGCATGATGCGGCATCTGCAGGCGCCAGGTAGTGCCTGCTTAG 60
|||||

Qy 189 ATGTCAAGCTGAAAAAACAAGAGAGACTGTGTGTGCTTGGGGAGATGTAATCATTC 248
|||||

Db 61 ATGTCAAGCTGAAAAAACAAGAGAGACTGTGTGTGCTTGGGGAGATGTAATCATTC 120
|||||

Qy 249 CTTCCACAACCTGCTGATGCTCCCTGTGGTGAAGAACAACATCGCTGCTCTTGCCA 308
|||||

Db 121 CTTCCACAACCTGCTGATGCTCCCTGTGGTGAAGAACAACATCGCTGCTCTTGCCA 180
|||||

Qy 309 GCAGACTGGGGTGTCCAAAGATCGCAATAGAGTGTAGAGGCTTTAGCCCA 368
|||||

Db 181 GCAGACTGGGGTGTCCAAAGATCGCAATAGAGTGTAGAGGCTTTAGCCCA 240
|||||

Qy 369 GTTGTTCAGAGCCCTGGTGGATCTTGTATCCAGGCGCTCAAGGCTAGAACCTCA 428
|||||

Db 241 GTTGTTCAGAGCCCTGGTGGATCTTGTATCCAGGCGCTCAAGGCTAGAACCTCA 300
|||||

Qy 429 GGGATGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGGACTCATCAAGC 486
|||||

Db 301 GGGATGAATCTTCAATAGAGCCGATGATCTGTGTC--TTTGGACTCATCAAGC 360
|||||

Qy 487 CTGGT--AGCATTTGCACTTTATTTATGAAATCTCTGTGATTAAGAGTAAT 544
|||||

Db 361 CTGGTTTAGCATTTTTCAGATTTATCTTCAGAAATCTCTGGCATTAAGAGTAAT 420
|||||

Qy 545 TATTAAGTGGTCTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 604
|||||

Db 421 TATTNAGTGTCTCTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
|||||

Qy 605 ATAAAAAGGAAGAGCTCCAAATGATCAC--TTATATTTTACCATTTCTATACAC 663
|||||

Db 481 ATAAAAAGGAAGAGCTCCAAATGATCACCTTTATATTTACCATTTCTATACAC 540
|||||

Qy 664 AGCAGTGAAGAGCTTTC--GAGACTTTTCAGTCTTATGTTGATCACTTAAAAAG 721
|||||

Db 541 AGCAGTGAAGAGCTTTC--GAGACTTTTCAGTCTTATGTTGATCACTTAAAAAG 600
|||||

Qy 722 AATGTACAGTACAAATAAAGTCAGTTTAA 754
|||||

Db 601 AATGTACAGTACAAATAAAGTCAGTTTAA 633
|||||

RESULT 42
CD556579 946 bp mRNA linear EST 11-JUN-2003
LOCUS AGENCOURT 14477072 NIH_MGC 179 Homo sapiens cDNA clone
DEFINITION IMAGE:30392712 5', mRNA sequence.
ACCESSION CD556579
VERSION CD556579.1 GI:31582647
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 946)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: NDAM468 row: h column: 01
High quality sequence start: 22
High quality sequence stop: 606.
Location/Qualifiers
1..946
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30392712"
/tissue_type="Pituitary"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"

FEATURES
source

/clone lib="NIH_MGC_179"
/note="Organ: brain; Vector: PCMV-SPORTS.1; Site: 1; Score:
(destroyed); Site 2: Not; Library: site is oligo-dt primed and
directionally cloned (EcoRV site is destroyed upon cloning
) . Average insert size 1.1 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."
BASE COUNT 221 a 262 c 236 g 227 t
ORIGIN

Query Match 74.2%; Score 559.2; DB 14; Length 946;
Best Local Similarity 95.1%; Pred. No. 1.6e-146;
Matches 656; Conservative 0; Mismatches 18; Indels 16; Gaps 7;

QY 1 ATGGCCGACGTGGAGAGAGAGAACTGGCCCTGCTCTCACTCCGGAGCTCA 60
Db 75 ATGGCCGACGTGGAGAGAGAGAACTGGCCCTGCTCTCACTCCGGAGCTCA 134
QY 61 GGCTCCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 135 GGCTCCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 194
QY 121 TGGAGCTGGAG 180
Db 195 TGGAGCTGGAG 247
QY 181 TGTCTTAGATGTCAAGCTGAAACAAACAAAGAGAGAGAGAGAGAGAGAG 240
Db 248 TGTCTTAGATGTCAAGCTGAAACAAACAAAGAGAGAGAGAGAGAGAGAG 307
QY 241 AATCATCTCTTCCACAACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 308 AATCATCTCTTCCACAACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 367
QY 301 CTCTGACAG 360
Db 368 CTCTGACAG 427
QY 361 TTAGCCGACGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 428 TTAGCCGACGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 487
QY 421 ACACCTACAG 478
Db 488 ACACCTACAG 547
QY 479 AATCAAGCCTTGCTT--AGCATTTGCTGATTTATCTTCAGAAATTTCT 536
Db 548 AATCAAGCCTTGCTT--AGCATTTGCTGATTTATCTTCAGAAATTTCT 607
QY 537 AGATAATTATTAAAGGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 595
Db 608 AGATAATTATTAAAGGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 667
QY 596 AGAAGGTCTATTA--AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
Db 668 AGAAGGTCTATTA--AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 727
QY 652 TTTCTATTACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 681
Db 728 TTTCTATTACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 757

RESULT 43
LOCUS CB217926 622 bp mRNA linear EST 06-FEB-2003
DEFINITION NISC nb05g10.y1 COGENE 6E MAN Homo sapiens cDNA clone IMAGE:579514
5', mRNA sequence.
ACCESSION CB217926
VERSION CB217926
KEYWORDS EST, GI:28266118
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 622)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
CDNA Library Preparation:
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plates: LLN12898 row: M column: 19
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1. 622

FEATURES

source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:579514"
/tissue_type="mandible, pooled"
/dev_stage="embryo, 6 weeks postconception"
/lab_host="DH10B"
/clone_lib="COGENE 6E MAN"
/note="Vector: pAMP1; cDNA primed using oligo-dt primer,
directionally cloned into UDO sites of pAMP1. Size
selected for insert sizes ranging from 0.2-2.0 kb.
Normalized to Cots. Primary library, non-amplified.
Library constructed by M. Lovett. For more information on
this library, please contact R. Tidwell (Washington
University) or visit the COGENE website at
http://hg.wustl.edu/COGENE/."

BASE COUNT

144 a 147 c 182 g 149 t

Query Match 74.1%; Score 558.8; DB 14; Length 622;
Best Local Similarity 98.2%; Pred. No. 1.7e-146;
Matches 587; Conservative 0; Mismatches 7; Indels 4; Gaps 2;

QY 1 ATGGCCGACGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 25 ATGGCCGACGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 84
QY 61 GGCTCCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 85 GGCTCCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 121 TGGAGCTGGAG 180
Db 145 TGGAGCTGGAG 204
QY 181 TGTCTTAGATGTCAAGCTGAAACAAACAAAGAGAGAGAGAGAGAGAGAG 240
Db 205 TGTCTTAGATGTCAAGCTGAAACAAACAAAGAGAGAGAGAGAGAGAGAG 264
QY 241 AATCATCTCTTCCACAACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 265 AATCATCTCTTCCACAACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 324
QY 301 CTCTGACAG 360
Db 325 CTCTGACAG 384
QY 361 TTAGCCGACGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 385 TTAGCCGACGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 444
QY 421 ACACCTACAG 478
Db 445 ACACCTACAG 504

Db	Accession	Definition	Version	Keywords	Source	Organism	Reference Authors Title	Journal	Medline	PubMed	Comment	Features
Oy	479	ATCAAGCGCTGGT-- AGACTTTGACGTTTATACCTCAAAATTCCTGTATTAGA	536									
Db	505	ATCAAGCGCTGGTTACATTTTGACAGTTTATCTTCAAAATTCCTGTATTAGA	564									
Oy	537	AGATAATTTATTAAGGTGTGCTCTTCTTCTGTGTGTGTGTGTGGACACAGCT	594									
Db	565	AGATAATTTATTAAGGTGTGCTCTTCTTCTGTGTGTGTGTGTGGACACAGCT	622									
RESULT 44												
LOCUS	BUT30720/c											
DEFINITION	UI-E-CII-1-8fo-o-15-0-UI.s1 UI-E-CII Homo sapiens cDNA clone	656 bp	linear	EST 09-OCT-2002								
ACCESSION	UI-E-CII-8fo-o-15-0-UI.s1											
VERSION	UI-E-CII-8fo-o-15-0-UI 3', mRNA sequence.											
KEYWORDS	EST.											
SOURCE	BUT30720.1 GI:23654893											
ORGANISM	Homo sapiens (human)											
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
AUTHORS	Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.											
TITLE	1 (bases 1 to 656)											
	Bonaldi,M.F., Lennon,G. and Soares,M.B.											
	Normalization and subtraction: two approaches to facilitate gene											
	discovery											
	Genome Res. 6 (9), 791-806 (1996)											
JOURNAL	Genome Res. 6 (9), 791-806 (1996)											
MEDLINE	97044477											
PUBMED	8889548											
COMMENT	Contact: Soares, MB											
	Coordinated Laboratory for Computational Genomics											
	University of Iowa											
	375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA											
	Tel: 319 335 8250											
	Fax: 319 335 9565											
	Email: bento-soares@iowa.edu											
	Tissue Procurement: Dr. Gregg Hageman											
	cDNA library preparation: Dr. M. Bento Soares, University of Iowa											
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa											
	Clone Distribution: Researchers may obtain clones from Research											
	Genetics (www.resgen.com).											
	Seg primer: M13 FORWARD											
	POLYA=Yes.											
	Location/Qualifiers											
	1..656											
	/organism="Homo sapiens"											
	/mol_type="mRNA"											
	/db_xref="taxon:9606"											
	/clone="UI-E-CII-8fo-o-15-0-UI"											
	/tissue_type="RPE and Choroid"											
	/dev_stage="adult"											
	/lab_host="DH10b (life Technologies) (T1 phage resistant)"											
	/clone_lib="UI-E-CII"											

	BASE COUNT	175 a	157 c	125 g	198 t	1 others
ORIGIN						
Query Match	74.0%; Score 558.2; DB 13;					Length 656;
Best Local Similarity	96.7%; Pred. No. 2.ee-146;					
Matches 613; Conservative	0; Mismatches 14; Indels 7; Gaps 4;					
OY	128	GGGACGTGGAGTGCATTCGCCCATCTCGCAGGGTGCCAGSTGATGATGCTTCTTTA	187			
Dd	656	GGAACGTGGAGTGCATTCGCCCATCTCGCAGGGTGCCAGSTGATGATGCTTCTTTA	597			
OY	188	GATGTCAAGCTGAAAACAACAAGAAGACCTGTGTGGGTCTGGGGGAGATGTAATCAT	247			
Dd	596	GATGTCAAGCTGAAAACAACAAGAAGACCTGTGTGGGTCTGGGGGAGATGTAATCAT	537			
OY	248	CCTTCACAACCTGTGCATGTCCCTGTGGGTGAAACGAAACAATCGCTGCCCTCTGCC	307			
Dd	536	CCTTCACAACCTGTGCATGTCCCTGTGGGTGAAACGAAACAATCGCTGCCCTCTGCC	477			
OY	308	AGCAGGACTGGGTGCTCCAAAGAATTGGCAAATCAGAGTGGTTAAGAGGCTTTTAGGCG	367			
Dd	476	AGCAGGACTGGGTGCTCCAAAGAATTGGCAAATCAGAGTGGTTAAGAGGCTTTTAGGCG	417			
OY	368	AGTTGTCAAGCCCTGGTGGATCTTGAAATCCAGTCCCTACAAAGGCTAGAACACTAC	427			
Dd	416	AGTTGTCAAGCCCTGGTGGATCTTGAAATCCAGTCCCTACAAAGGCTAGAACACTAC	357			
OY	428	AGGGGATGAATCTTCCAATATGAGGCCGATGATCTGTGTC--TTTGACTCATCAAG	485			
Dd	356	AGGGGATGAATCTTCCAATATGAGGCCGATGATCTGTGTC--TTTGACTCATCAAG	297			
OY	486	CCTTGCTT--AGCAATTGTGCTTATPACTCACAAATCTCTGTGATTAAGAAGATAAT	543			
Dd	236	CCTTGCTTATGACATTTTGTCTGATTTTATCTTCAAAATCTCTGTGATTAAGAAGATAAT	237			
OY	544	TATATTAAGTGTCTTCTTCTACCCTGTGTGTGTGTGCGGACAACAGCTTAAGAAGTC	603			
Dd	236	TATATTAAGTGTCTTCTTCTACCCTGTGTGTGTGTGCGGACAACAGCTTAAGAAGTC	177			
OY	604	TATAAAAAGAAAGAGTCCAAATFTGAATCACC--TTATATTTACCATTCTATACAA	662			
Dd	176	TATAAAAAGAAAGAGTCCAAATFTGAATCACC--TTATATTTACCATTCTATACAA	117			
OY	663	CAGCAGATGSAAGCAGATTTC--GAGACTTTTTGANTGCTTAAGTTGATCACTTAACAAA	720			
Dd	116	CAGCAGATGSAAGCAGATTTC--GAGACTTTTTGANTGCTTAAGTTGATCACTTAACAAA	57			
OY	721	GATGTTCAGTAAACAATTAAGTCCATTTAA	754			
Dd	56	GATGTTCAGTAAACAATTAAGTCCATTTAA	23			
RESULT 45						
BFOJ3587		595 bp	mRNA	linear	EST 20-OCT-2000	
LOCUS	601453564F1 NIH_MGC_66 Homo sapiens CDNA clone IMAGE:3857511 5'					
DEFINITION	mRNA sequence.					
ACCESSION	BFOJ3587					
VERSION	BFOJ3587.1 GI:10741299					
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 595) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: cgabos-e@mail.nih.gov Tissue Procurement: DClD/DTP CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL).					

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM587 row: 0 column: 16
 High quality sequence stop: 595.

FEATURES

Source

1. 595
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3857511"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH MGC 66"
 /note="Organ: ovary; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."
 BASE COUNT 153 a 126 c 165 g 151 t
 ORIGIN

Query Match 73.7%; Score 555.4; DB 10; Length 595;
 Best Local Similarity 98.2%; Pred. No. 1.6e-145;
 Matches 584; Conservative 0; Mismatches 6; Indels 5; Gaps 2;

46 CACTCCGGAGCTCAGGCTCCAAAGTGGGAGGAGGAGAGATGTTCCCTCCCAAGAGTGG 105
 1 CACTCCGGAGCTCAGGCTCCAAAGTGGGAGGAGGAGAGATGTTCCCTCCCAAGAGTGG 60
 106 AACGGGTGCGCATGTGGAGCTGGAGCTGGAGTGCATGCTGCCCATCTGACAGGCTC 165
 61 AACGGGTGCGCATGTGGAGCTGGAGCTGGAGTGCATGCTGCCCATCTGACAGGCTC 120
 166 CAGGTGATGATGCTCTCTTAAATGTCAGAGCTGAACCAACAGAGAGAGCTGTGGT 225
 121 CAGGTGATGATGCTCTCTTAAATGTCAGAGCTGAACCAACAGAGAGAGCTGTGGT 180
 226 GTCGCGGAGAGATGATATCATTCCTCCCAACCTGCTGCATGCTCCCTGTGGTGAACAG 285
 181 GTCGCGGAGAGATGATATCATTCCTCCCAACCTGCTGCATGCTCCCTGTGGTGAACAG 240
 286 AACATGCTGCTGCTCTCTCCAGCAGAGCTGGGTGCTCCAAAGATCGCAATGAGAG 345
 241 AACATGCTGCTGCTCTCTCCAGCAGAGCTGGGTGCTCCAAAGATCGCAATGAGAG 300
 346 TGGTTGAAGGCTTCTTAGCGCAGTTGTTCAAGCCTGTGATCTTGAATCAAGTC 405
 301 TGGTTGAAGGCTTCTTAGCGCAGTTGTTCAAGCCTGTGATCTTGAATCAAGTC 360
 406 CCTACAAAGGCTTGAACACTACAGGGAGATGAATTTCTCAATAGAGCCGATGATCTGT 465
 361 CCTACAAAGGCTTGAACACTACAGGGAGATGAATTTCTCAATAGAGCCGATGATCTGT 420
 466 GGTCTCTTGGACTCATCAAGCCTTGGTT--AGCATTTGTCAAGTTTATCTTCAGAAA 520
 421 GGTCTCTTGGACTCATCAAGCCTTGGTTTGTAGCATTTTGTCAAGTTTATCTTCAGAAA 480
 521 TTTCTGTGTATTAAGAGATTAATTTAATAAGGTGCTCTTCTTACTCTGTGTGTGTG 580
 481 TTTCTGTGTATTAAGAGATTAATTTAATAAGGTGCTCTTCTTACTCTGTGTGTGTG 540
 581 TGGCGACACAGCTTGAAGATGCTATTAATAAGAGAGAGCTCCAATTGAATCA 635
 541 TGGCGACACAGCTTGAAGATGCTATTAATAAGAGAGAGCTCCAATTGAATCA 595

RESULT 46
 BG026092 781 bp mRNA linear EST 24-JAN-2001
 LOCUS 602322288F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4387076 5',
 DEFINITION mRNA sequence.
 ACCESSION BG026092

VERSION BG026092.1 GI:12413350
 EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 781)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC

COMMENT

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM10070 row: p column: 21
 High quality sequence stop: 777.
 Location/Qualifiers

FEATURES

Source

1. 781
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4387076"
 /tissue_type="osteosarcoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH MGC 86"
 /note="Organ: bone; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.53 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC library."
 BASE COUNT 216 a 163 c 210 g 192 t
 ORIGIN

Query Match 73.5%; Score 554.2; DB 10; Length 781;
 Best Local Similarity 94.8%; Pred. No. 3.8e-145;
 Matches 693; Conservative 0; Mismatches 23; Indels 15; Gaps 11;

24 GGAACCTGCGGCTGCTGCTCCATCCGCGGAGCTGAGCTCCAACTCGGGAGGCGACAA 83
 26 GGAACCTGCGGCTGCTGCTCCATCCGCGGAGCTGAGCTCCAACTCGGGAGGCGACAA 85
 84 GATGTTCTCCCTCAAGAGTGAACGCGGTGCGCATGTGAGCTGGAGCTGGAGTGCAG 143
 86 GATGTTCTCCCTCAAGAGTGAACGCGGTGCGCATGTGAGCTGGAGCTGGAGTGCAG 144
 144 TACGTGCGGCATCTGAGGGTCCAGGTGATGATGCTCTTGAATGTCAGAGCTGAAGA 203
 145 TACGTGCGGCATCTGAGGGTCCAGGTGATGATGCTCTTGAATGTCAGAGCTGAAGA 204
 204 CAAACAGAGAGCTGTGTTGTGCTGCGGAGAAATGATCATTCCTCTCCAAACGCGG 263
 205 CAAACAGAGAGCTGTGTTGTGCTGCGGAGAAATGATCATTCCTCTCCAAACGCGG 263
 264 CATGTCCTGTGGGTGAACAGAACATGCTGCTCTCTCCAGCAGAGCTGGGTGTG 323
 264 CATGTCCTGTGGGTGAACAGAACATGCTGCTCTCTCCAGCAGAGCTGGGTGTG 323
 324 CCAAGAGATCGGCAATGAGAGTGTAGAGGCTTCTTAGGCGAGTTGTCAGAGCCTT 383
 324 CCAAGAGATCGGCAATGAGAGTGTAGAGGCTTCTTAGGCGAGTTGTCAGAGCCTT 382
 384 GGTGATCTTGAATCAAGTGCCTTACAAAGGCTTGAACACTACAGGGAGATGAATCTTC 443
 383 GGTGATCTTGAATCAAGTGCCTTACAAAGGCTTGAACACTACAGGGAGATGAATCTTC 442
 444 AATAGAGAGCCGATGATCTGTGTC--TTTGAATCATCAAGCCTTGTGTT--AGCATT 499
 443 AATAGAGAGCCGATGATCTGTGTCCTTTGGAGCTCATCAAGCCTGTGTTTGAATTT 502

[illegible]

RESULT	47
AM001261/c	
LOCUS	AM001261
DEFINITION	wb59g7.v1 Soares Dieckreife col-on NHCD Homo sapiens cDNA clone IMAGE:2521212 3' similar to WP:R10A10.2 CE12670 ZINC FINGER PROTEIN
Accession	.. mRNA sequence.

ACCESSION	AW001261	GI:5848177
VERSION	AW001261.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
Tumor Gene Index

Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 456.

FEATURES	Location/Qualifiers
source	1. .670

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2521212"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"

```

```

/!ad,host="PH10B(phage-resistant)"
/clone.lib="Stares Dieckgraebe_colon_NHCD"
/notes="Organ: colon; Vector: pPT3D-Pac (Pharmacia) with
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAAGTGGAGGCGGCGCCCTCTTTTATTTTATTTT
3')],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
and Eco RI sites of the modified pT733 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraebe (Washington University,
dieck@im.wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento

```

BASE COUNT
ORIGIN

Query Match	73.4%	Score 553.6;	DB 9;	Length 670;
Best Local Similarity	95.0%;	Pred. No. 5.3e-145;		
Matches 640;	Conservative 0;	Mismatches 19;	Indels 15;	Gaps 6

QY	88	TTCTCCCTCCAAAGAGTGAAACCGGTTGGCCATGTGAGCTGGGACCTGGAAGTGCATACG	147
Db	670	TTCTCCATCAAGAGTGAAACCGGGTGCCCATGTGAAGCTGGGACGTGAGTCCATACG	611
QY	148	TGGCGCATCTGAGAGGTCCAGGTATGATGCTCTCTTAAGATGTAAAGCTGAAACAA	207
Db	610	TGGCGCATCTGAGAGTGTCA-----GATGGGTGTCCTTAAGTGTCAACTGACAAAC-AA	559
QY	208	CAAGAGCACTGTGTGTGGTCTGGGGAGATGATTAATCATCTCTTCACAACTCTGCATG	267
Db	558	CAAGAGCACTGTGTGTGGTCTGGGGAGATGATTAATCATCTCTTCACAACTCTGCATG	499
QY	268	TCCCTGTGGGTGAAAACAGAACATTCGCTGCCTCTCTTCGCACGACAGACTGGGGTGTCAA	327
Db	498	TCCCTGTGGGTGAAAACAGAACATTCGCTGCCTCTCTTCGCACAGACTGGGGTGTCCA	439
QY	328	AGAACTGGCAAAATGAGAGTGGTTAGAAGCTTTCTTAGGCGCATGTGTTCAAGCCCTGGTG	387
Db	428	AGAACTGGCAAAATGAGAGTGGTTAGAAGCTTTCTTAGGCGCATGTGTTCAAGCCCTGGTG	379
QY	388	GATCTTGTAAATCAAGTGGCCCTACAAAGGCTAGAACACTACAGGGAGTGAATTCCTCAAAAT	447
Db	378	GATCTTGTAAATCAAGTGGCCCTACAAAGGCTAGAACACTACAGGGAGTGAATTCCTCAAAAT	319
QY	448	AGGAGCCGATGGATCTGTGTGCTTT--GGACTCATCAAAAGCTTGGTT--AGCATTTGTTC	503
Db	318	AGGAGCCGATGGATCTGTGTGCTTTGGGACATCAAAAGCCTTGGTTAAGCATTTTGTTC	259
QY	504	AGTTTATCTTCAGAAATTTCTCTGTGATTAAGAAATATTAATTAATAAGTGTCTCTTC	563
Db	258	AGTTTATCTTCAGAAATTTCTCTGTGATTAAGAAATATTAATTAATAAGTGTCTCTTC	199
QY	564	TACCTCTGTGTGTGTGTGGCGCACACAGCTTAGAAGTGTCTATAAAAAAGAAAGAGCTC	623
Db	198	TACCTCTGTGTGTGTGTGTGGCGCACACAGCTTAGAAGTGTCTATAAAAAAGAAAGAGCTC	139
QY	624	CAAAATGGAATCAACC-TTAAATTTAACCATTTCTATAACACAGGAGAGTGAAGCAGTTTC	682
Db	138	CAAAATGGAATCAACTTTATTAATTTTATTAATTTCTATAACACAGGAGAGTGAAGCAGTTTC	79
QY	683	--GAGACTTTTTCGATGCTTAATGTGTGATCAGTTAAAAAAGATGTTCACGTAAACAATA	740
Db	78	AGAGAACTTTTTCGATGCTTAATGTGTGATCAGTTAAAAAAGATGTTCACGTAAACAATA	19
QY	741	AAAGTCAGTTTAAA 754	
Db	18	AAAGTACAGTTTAAA 5	

RESULT 48					
LOCUS	Bg111145				
DEFINITION	Bg111145	793 bp	mRNA	linear	EST 30-JAN-2001
ACCESION	602281555P1	NIH_MGC_86	Homo sapiens	CDNA clone	IMAGE:4365332 5',
VERSION	Bg111145				
KEYWORDS	Bg111145.1	GI:12604651			
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 793)				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: gsabbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc.				

Query Match 73.3%; Score 552.8; DB 12; Length 654;
 Best Local Similarity 97.0%; Pred. No. 8.8e-145;
 Matches 607; Conservative 0; Mismatches 12; Indels 7; Gaps 4;

```

QY 136 GAGTCCGATACGTGGCCATCTGCAAGGTCAGGTCATGATGATGCTGTCTTAAATGATGTA 195
DB 648 GAGGCGGATACGTGGCCATCTGCAAGGTCAGGTCATGATGATGCTGTCTTAAATGATGTA 589
QY 196 GCTGAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 255
DB 588 GCTGAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
QY 256 AACTGTGATGTCCTCTGTGGTGAACAGAACAAATGCTGCTCTCTCTGACAGAGAGAG 315
DB 528 AACTGTGATGTCCTCTGTGGTGAACAGAACAAATGCTGCTCTCTCTGACAGAGAGAG 469
QY 316 TGGGTGTCTCAAGAAATGCGCAATGAGAGTGTGTAAGAGAGAGAGAGAGAGAGAGAGAG 375
DB 468 TGGGTGTCTCAAGAAATGCGCAATGAGAGTGTGTAAGAGAGAGAGAGAGAGAGAGAGAG 409
QY 376 AGAGCCCTGTGATGTCCTGTGAATCCAGTCCCTACAAAGGCTAGAGAGAGAGAGAGAG 435
DB 408 AGAGCCCTGTGATGTCCTGTGAATCCAGTCCCTACAAAGGCTAGAGAGAGAGAGAGAG 349
QY 436 AATTTCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 493
DB 348 AATTTCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 289
QY 494 --AGCATTTGTGATGTCCTGTGAATCCAGTCCCTACAAAGGCTAGAGAGAGAGAGAG 551
DB 288 TAGCATTTGTGATGTCCTGTGAATCCAGTCCCTACAAAGGCTAGAGAGAGAGAGAGAG 229
QY 552 GGTGTCTCTCTCACTCTGTGGTGTGTGCGGACACAGCTTAGAGTGTATATAAA 611
DB 228 GGTGTCTCTCTCACTCTGTGGTGTGTGCGGACACAGCTTAGAGTGTATATAAA 169
QY 612 AGGAAAGAGCTCCAAATGATGATC--TTATAATTTTACCATTCTATACACAGGAGCT 670
DB 168 AGGAAAGAGCTCCAAATGATGATC--TTATAATTTTACCATTCTATACACAGGAGCT 109
QY 671 GGAAGAGATTC--GAGACTTTTGTGATGTCATGTCATGTCATGTCATGTCATGTCAT 728
DB 108 GGAAGAGATTC--GAGACTTTTGTGATGTCATGTCATGTCATGTCATGTCATGTCAT 49
QY 729 CAGTAACAATAAAGTCAAGTTTAA 754
DB 48 CAGTAACAATAAAGTCAAGTTTAA 23

```

RESULT 50
 BG718227 715 bp mRNA linear EST 08-MAY-2001
 LOCUS 602696194F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4828198 5',
 DEFINITION mRNA sequence.
 ACCESSION BG718227
 VERSION BG718227.1 GI:13997414
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 715)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: gspbs-remail.nih.gov
 Tissue Procurement: Miklos Palakovic, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

Found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LHAM0745 row: 1 column: 23
 High quality sequence stop: 714.
 Location/Qualifiers
 1. 715

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4828198"
/lab_host="DH10B"
/clone_lib="NIH MGC 97"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTATTTTATTTTAA-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 151 a 184 c 217 g 163 t
ORIGIN

```

Query Match 72.4%; Score 546.2; DB 10; Length 715;
 Best Local Similarity 98.0%; Pred. No. 6.7e-143;
 Matches 575; Conservative 0; Mismatches 8; Indels 4; Gaps 2;

```

QY 1 ATGSCCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
DB 129 ATGSCCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 188
QY 61 GGTCTCAAGTGGGAGGCGACAGATGTTCTCTCTCAAGAGTGAACCGGCTGGCCATG 120
DB 189 GGTCTCAAGTGGGAGGCGACAGATGTTCTCTCTCAAGAGTGAACCGGCTGGCCATG 248
QY 121 TGGAGCTGGGAGCGGAGTGGATGATGCGCCATCTGCAAGGTCACAGTGTATGATGCC 180
DB 249 TGGAGCTGGGAGCGGAGTGGATGATGCGCCATCTGCAAGGTCACAGTGTATGATGCC 308
QY 181 TGTCTTAGATGTCAGCTGAGAAACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 309 TGTCTTAGATGTCAGCTGAGAAACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 368
QY 241 AATCATTCCTCCCAACTGCTGATGTCCTCTGTGGTGAACAGAGAAATCGTGCCT 300
DB 369 AATCATTCCTCCCAACTGCTGATGTCCTCTGTGGTGAACAGAGAAATCGTGCCT 428
QY 301 CTCTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 429 CTCTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 488
QY 361 TTAGGCGAGTGTTCAGAGCCCTGTGATGTCATGTCATGTCATGTCATGTCATGTCAT 420
DB 489 TTAGGCGAGTGTTCAGAGCCCTGTGATGTCATGTCATGTCATGTCATGTCATGTCAT 548
QY 421 ACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478
DB 549 ACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 608
QY 479 ATCAAGGCTTGTT--AGCATTTGTGATGTCATGTCATGTCATGTCATGTCATGTCAT 536
DB 609 ATCAAGGCTTGTT--AGCATTTGTGATGTCATGTCATGTCATGTCATGTCATGTCAT 668
QY 537 AGATATTTTATTAAGGTGTGTCCTTCTGATGTCATGTCATGTCATGTCATGTCATGTC 583
DB 669 AGATATTTTATTAAGGTGTGTCCTTCTGATGTCATGTCATGTCATGTCATGTCATGTC 715

```

Search completed: November 7, 2003, 15:15:50
 Job time: 1912.92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 11:11:37 ; Search time 823.044 Seconds
(without alignments)
6163.455 Million cell updates/sec

Title: US-09-509-779-3_COPY_141_264
Perfect score: 124
Sequence: 1 CGATACGTCGCGCATCTGCA.....ATTCTTCACACATGCTGC 124

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : GenBank1.*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sbs:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_mam:*
36: em_htg_rtd:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	124	100.0	342	6	AX282592	Sequence
2	124	100.0	342	9	AF142060	Homo sapi
3	124	100.0	342	9	BT007348	Homo sapi
4	124	100.0	754	6	BD096967	SAG:apopt
5	124	100.0	754	6	BD096989	SAG:apopt
6	124	100.0	754	6	BD096994	SAG:apopt
7	124	100.0	754	6	BD096995	SAG:apopt
8	124	100.0	754	6	AF092878	Homo sapi
9	124	100.0	816	9	BC008627	Homo sapi
10	124	100.0	822	9	BC005966	Homo sapi
11	124	100.0	836	9	AF164679	Homo sapi
12	121	97.6	754	6	BD096988	SAG:apopt
13	116	93.5	754	6	BD096981	SAG:apopt
14	112	90.3	754	6	BD096987	SAG:apopt
15	107	86.3	754	6	BD096982	SAG:apopt
16	107	86.3	754	6	BD096983	SAG:apopt
17	103	83.1	754	6	BD096986	SAG:apopt
18	103	83.1	754	6	BD096990	SAG:apopt
19	97	78.2	754	6	BD096985	SAG:apopt
20	90	72.6	747	6	BD096973	SAG:apopt
21	89	71.8	674	9	AF312226	Homo sapi
22	83	66.9	754	6	BD096984	SAG:apopt
23	76	61.3	754	6	BD096992	SAG:apopt
24	74	59.7	754	6	BD096991	SAG:apopt
25	74	59.7	754	6	BD096993	SAG:apopt
26	49	39.5	35638	9	AC112771	Homo sapi
27	49	39.5	125041	2	AC068693	Homo sapi
28	49	39.5	238330	2	AC106176	Rattus no
29	49	39.5	245476	2	AC098496	Rattus no
30	48	38.7	187353	2	AC118192	Mus muscu
31	43	34.7	324	6	BD077705	5'EST of
32	43	34.7	59155	2	AC104983	Homo sapi
33	43	34.7	163521	2	AC108040	Homo sapi
34	43	34.7	164500	2	AC092418	Homo sapi
35	42	33.9	706	6	BD096974	SAG:apopt
36	37	29.8	254832	2	AC095698	Rattus no
37	36	29.0	120515	10	AF929565	Mouse DNA
38	35	28.2	439	9	HA323208	Homo sapi
39	35	28.2	264606	2	AC137264	Rattus no
40	32	25.8	1125	10	BC011127	Mus muscu
41	32	25.8	1140	6	BD096966	SAG:apopt
42	32	25.8	1140	10	AF092877	Mus muscu
43	32	25.8	219370	11	AC123935	Mus muscu
44	31	25.0	612	11	BV071691	S212P6546
45	31	25.0	112321	2	AC103666	Mus muscu

ALIGNMENTS

RESULT 1	AX282592	342 bp	DNA	linear	PAT 02-NOV-2001
LOCUS	AX282592				
DEFINITION	Sequence 7 from Patent WO0175145.				
ACCESSION	AX282592				
VERSION	AX282592.1				
KEYWORDS	GI:16609675				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Isakani,S.D., Huang,J., Sheung,J. and Pray,T.R.				
AUTHORS	1				
TITLE	Ubiquitin ligase assay				
JOURNAL	Patent: WO 01/5145-A 7 11-OCT-2001;				

	Matches	124;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
OY	1	CGATACGTCGCCCATCTGCAGGGGTCAAGGTGATGATGCTTGTCTTAAGTGTCAA	60							
Dd	141	CGATACGTCGCCCATCTGCAGGGGTCAAGGTGATGATGCTTGTCTTAAGTGTCAA	200							
OY	61	AAACAAACAAGAGA CTGTGTTGTGGTCTGGGGAGAATGTA TCATTCC TCCACA	120							
Dd	201	AAACAAACAAGAGA CTGTGTTGTGGTCTGGGGAGAATGTA TCATTCC TCCACA	260							
OY	121	CTGC 124								
Dd	261	CTGC 264								

RESULT 4				
BD096967				
LOCUS	BD096967	754 bp	DNA	linear
DEFINITION	SAG:apoptosis sensitivity gene.			
ACCESSION	BD096967			
VERSION	BD096967.1	GI:22642555		
KEYWORDS	JP 2001526063-A/2.			
SOURCE	unidentified			
ORGANISM	unidentified			
REFERENCE	unclassified.			
AUTHORS	1 (bases 1 to 754)			
TITLE	Sun,Y.			
JOURNAL	SAG:apoptosis sensitivity gene			
	Patent: JP 2001526063-A 2			
	18-DEC-2001;			
	WARNER LAMBERT CO			
COMMENT	OS	Unidentified		

PF	15-DEC-1998 JP	2000525451	
PR	19-DEC-1997 US	60/068179, 11-SEP-1998 US	60/099840 PI
YI	SRIN		
PC	C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P17/02, A61P35/00,		
PC	A61P39/06,		
PC	A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC		
	, C12N5/10, C12Q1/68,		
PC	G01N33/50, G01N33/68, C12N15/00, A61K37/02, C12N5/00 CC		
Strandedness:	Double;		
CC	Topology: Linear;		
CC	/note = 'Human SAG'		
PH	Key	Location/Qualifiers	
FT	CDS	1..339	
FT	mat_peptide	1..339	
FT	misc_feature	1..754.	
FT	Location/Qualifiers		
	1..754		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32654"		
BASE COUNT	205 a	155 c	201 g 193 t
ORIGIN			

Query Match	Similarity	100.0%	Score 124	DB 6	Length 754
Best Local	Similarity	100.0%	Pred. No. 6.4e-66		
Matches	124	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY	1	CGATACGTGCGCCATCTGCAGGGGTCCAGGTGATGATGCTGCTCTTACATGTCACAGCTGA	60		
Db	141	CGATACGTGCGCCATCTGCAGGGGTCCAGGTGATGATGCTGCTCTTACATGTCACAGCTGA	200		
QY	61	AAACCAACAGAGGAGCTGTGTGTGTGCTCGGGAGAGATTAATCATTTCTTCCACAACCTG	120		
Db	201	AAACCAACAGAGGAGCTGTGTGTGTGCTCGGGAGAGATTAATCATTTCTTCCACAACCTG	260		
QY	121	CTGC	124		
Db	261	CTGC	264		

RESULT 5				
BD096989				
LOCUS	BD096989	754 bp	DNA	linear
DEFINITION	SAG:apoptosis sensitivity gene.			PAT 27-AUG-2002
ACCESSION	BD096989			
VERSION	BD096989.1 GI:22642577			
KEYWORDS	JF 2001526063-A/24.			
SOURCE	unidentified			
ORGANISM	unclassified			
REFERENCE	unclassified.			
AUTHORS	1. (bases 1 to 754)			
TITLE	Sun,Y.			
JOURNAL	SAG:apoptosis sensitivity gene			
	Patent : JP 2001526063-A 24 18-DEC-2001 ;			
	WARNER LAMBERT CO			
COMMENT	OS Unidentified			
	DS TD unclassified 1/24			

PR	19-DEC-1997	US	60/068179.11-SEP-1998	US	60/099840	PI
YI	SUN					
PC	C12N5/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,					
PC	A61P39/06,					
PC	A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21					PC
C12N5/10,C12Q1/68						
PC	G01N33/50,G01N33/68,C12N5/00,A61K37/02,C12N5/00					CC
Strandedness:	Double;					
CC	Topology:	Linear;				
CC	SAG:	apoptosis sensitivity gene				
PH	Key	Location/Qualifiers				
FT	CDS	1..339				
FT	mat_peptide	1..339				
	Location/Qualifiers					
	1..754					
	/organism="unidentified"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:32644"					
BASE COUNT	206 a	155 c	201 g	192 t		
ORIGIN						

	Query Match	Similarity	100.0%	Score 124	DB 6	Length 754
	Best Local	Similarity	100.0%	Pred. No. 6.4e-66		
	Matches	124	Conservative	0	Mismatches	0
					Indels	0
					Gaps	0
QY	1	CGATACGTCGCCCATCTGCAGGGTCCAGGTGATGATGCGCTGTCTAGATGTCACAGCTGA	60			
Db	141	CGATACGTCGCCCATCTGCAGGGTCCAGGTGATGATGCTGTCTTAGATGTCAACTGA	200			
QY	61	AAACCAACAAAGAGACTGTGTTGTGTCTGGGGAAGATGTATCATTTCTTCACAACCTG	120			
Db	201	AAACCAACAAAGAGACTGTGTTGTGTCTGGGGAAGATGTATCATTTCTTCACAACCTG	260			
QY	121	CTGC 124				
Db	261	CTGC 264				

LOCUS	BD096994	754 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	SAC:apoptosis sensitivity gene.				
ACCESSION	BD096994				
VERSION	BD096994.1	GI:22642582			
KEYWORDS	JP 2001526063-A/29.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	unclassified.				
AUTHORS	1 (bases 1 to 754)				
TITLE	Sun, Y.				
JOURNAL	SAC:apoptosis sensitivity gene				
COMMENT	Patient: JP 2001526063-A 25 18-DEC-2001;				
	WARNER LAMBERT CO				
	OS Unidentified				

FEATURES	source	location/Qualifiers
Query Match	100.0%;	Score 124; DB 6; Length 754;
Best Local Similarity	100.0%;	Pred. No. 6,4e-66;
Matches 124; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
BASE COUNT	206 a	155 c 201 g 192 t
ORIGIN		
Oy	1	CGATACGTCGCCCATCTGCAAGGCTCCAGAGTGATGATGACCTGCTTGTAGATGTCAAGCTGA 60
Db	141	CGATACGTCGCCCATCTGCAAGGCTCCAGAGTGATGATGACCTGCTTGTAGATGTCAAGCTGA 200
Oy	61	AAACAAACMAGAGACATCTGTGTGTGTGTGGGGGAATGTATATCTTCACCAACTG 120
Db	201	AAACAAACMAGAGACATCTGTGTGTGTGTGGGGGAATGTATATCTTCACCAACTG 260
Oy	121	CTGC 124
Db	261	CTGC 264
RESULT 8		
AF092878		754 bp mRNA linear PRI 24-JUL-2001
LOCUS		Homo sapiens zinc RING finger protein SAG mRNA, complete cds.
DEFINITION		AF092878
VERSION		AF092878.1 GI:4588033
KEYWORDS		
SOURCE		Homo sapiens (human)
ORGANISM		Homo sapiens
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS		1 (bases 1 to 754)
TITLE		Duan,H., Wang,Y., Aviram,M., Swaroop,M., Loo,J.A., Bian,J., Tian,Y., Mueller,T., Bisgaier,C.L. and Sun,Y. SAG, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents Mol. Cell. Biol. 19 (4), 3145-3155 (1999)
JOURNAL		
MEDLINE		99182502
PUBMED		10082581
REFERENCE		2 (bases 1 to 754)
AUTHORS		Sun,Y.
TITLE		Alterations of SAG mRNA in human cancer cell lines: requirement for the RING finger domain for apoptosis protection
JOURNAL		
MEDLINE		99435944
PUBMED		10506102
REFERENCE		3 (bases 1 to 754)
AUTHORS		Swaroop,M., Bian,J., Aviram,M., Duan,H., Bisgaier,C.L., Loo,J.A. and Sun,Y.
TITLE		Expression, purification, and biochemical characterization of SAG, a RING finger redox-sensitive protein
JOURNAL		
REFERENCE		4 (bases 1 to 754)
AUTHORS		Swaroop,M., Wang,Y., Miller,P., Duan,H., Jackoe,T., Madore,S.J. and Sun,Y.
TITLE		Yeast homolog of human SAG/ROC2/Rbx2/Hrt2 is essential for cell growth, but not for germination: chip profiling implicates its role in cell cycle regulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		
MEDLINE		203059864
PUBMED		10851089
REFERENCE		5 (bases 1 to 754)
AUTHORS		Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
TITLE		Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
JOURNAL		

MEDLINE	PUBMED	11252847
REFERENCE	6	(bases 1 to 754)
AUTHORS	Sun, Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-SEP-1998) Department of Molecular Biology, Parke-Davis, 2800 Plymouth Rd., Ann Arbor, MI 48105, USA	
FEATURES	Location/Qualifiers	
SOURCE	1..754	/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/chromosome="3"
		/map="3q22-q24"
		/cell_line="HeLa, D98/AH-2, HPRT-"
	1..342	
CDS	/function="growth promotion"	
	/note="redox sensitive; metal binding; expression protects cells from apoptosis induced by redox compounds"	
	/codon_start=1	
	/product="zinc RING finger protein SAG"	
	/protein_id="AAD25962.1"	
	/db_xref="GI:4588034"	
	/translation="MADVDEDEFCTALASHSGSSGSGDKMESLKKNNAVMAMSDP VECTCALCRQVWDACLRCAENKQEDCVVWGBCNHSFHNCCMSLWVKONNCPIC QQDWVDVRIGK"	
BASE COUNT	205 A	155 C 201 G 193 T
ORIGIN		
Query Match	100.0%; Score 124;	DB 9; Length 754;
Best Local Similarity	100.0%; Pred. No. 6,4e-66;	
Matches 124; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
Dy	1	CGATACGTGGCCCCATTCTCAGAGTCCAGGTATGATGATCCTGTCTTAGATGTCAAAGTCGA 60
	141	CGATACGTGGCCCCATTCTCAGAGTCCAGGTATGATGATCCTGTCTTAGATGTCAAAGTCGA 200
Dy	61	AAACAACAAGAGAAGCATGTGTTGTGTGTGGGAGAAATGTAATCATTCCTTCCAACAAGTG 120
	201	AAACAACAAGAGAAGCATGTGTTGTGTGTGGGAGAAATGTAATCATTCCTTCCAACAAGTG 260
Dy	121 CTGC 124	
Dd	261 CTGC 264	
RESULT 9		
LOCUS	BC008627	816 bp mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, ring finger protein 7, clone MGC:117274 IMAGE:4177613,	
VERSION	mRNA, complete cds.	
KEYWORDS	BC008627	
ACCSSION	BC008627.1	GI:14250388
ORGANISM	MGC.	
SOURCE	Homo sapiens (human)	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 816)	
AUTHORS	Strausberg, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: David N. Louis, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing By: Baylor College of Medicine Human Genome Sequencing Center	

Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAC Plates: 12 Row: h Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.
Location/Qualifiers
1. 816
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LOCUSID:9616"
/db_xref="taxon:9606"
/clone="MGC:17274 IMAGE:4177613"
/tissue_type="Brain, anaplastic oligodendroglioma with 1p/19q loss"
/clone_id="NCI CGAP_Brn67"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
37. 378
/codon_start=1
/product="ring finger protein 7"
/protein_id="AAH08627.1"
/db_xref="GI:14250389"
/translation="MAVDEGEETCALASHSGSKSGDKMFSLLKNNAAVAMWSMDVECTCAICRVQVMDACLRCQAEKQEDCVVAGECNHSEFNHCMSLWVKONNRCPCLCQDMVVRIRK"

BASE COUNT 226 a 178 c 213 g 199 t

ORIGIN

Query Match 100.0%; Score 124; DB 9; Length 816;
Best Local Similarity 100.0%; Pred. No. 6,5e-66;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAATCGTGGCCATTCGACGGCTGCAGGATGATGGATGCGCTGTTGATGTCAAGCTGA 60
|||
Db 177 CGATACGTGGCCATTCGACGGCTGCAGGATGATGGATGCGCTGTTGATGTCAAGCTGA 236

QY 61 AAACAACAAGAGGACTGTGTGTGTGTGTGGGAGATGTAAATCATTCCTTCCACAACCTG 120
|||
Db 237 AAACAACAAGAGGACTGTGTGTGTGTGTGGGAGATGTAAATCATTCCTTCCACAACCTG 296

QY 121 CTGC 124
|||
Db 297 CTGC 300

RESULT 10 822 bp mRNA linear PRI 12-JUN-2001
BC005966
LOCUS BC005966
DEFINITION Homo sapiens, ring finger protein 7, clone MGC:14618 IMAGE:4069078,
mRNA, complete cds.
ACCESSION BC005966
VERSION BC005966.1 GI:13543635
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIN-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINU)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-sngc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINU at: <http://image.lnl.gov>
Series: IRAL Plate: 21 Row: 0 Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, similarity but not identity to protein.
Location/Qualifiers

FEATURES

SOURCE

CDS

1..822
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/clone="MGC:14618 IMAGE:4069078"
/issue_type="Kidney, hypernephroma"
/clone_lib="NIH MGC_58"
/lab_host="DH10B"
/note="Vector: pDNR-LTB"
27..368
/codon_start=1
/product="ring finger protein 7"
/protein_id="AAH05966.1"
/db_xref="GI:13543636"
/translation="MADVEDGEETCALASHSSGSGSGDMPFLKKNNAVAMMSWD
VECDTCAICRYVMDACLRCAENKQEDCVVWGECHNSFHNCMSLWVKNRCPIC
QDDWVVRIGK"

BASE COUNT

240 a 175 c 210 g 197 t

ORIGIN

Query Match 100.0%; Score 124; DB 9; Length 822;
Best Local Similarity 100.0%; Pred. No. 6.5e-66;

Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTCGCCATCTGCAGGCTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 60
|||
Db 167 CGATACGTCGCCATCTGCAGGCTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 226
|||
QY 61 AAACAAACAGAGAGACTGTGTGTGCTGCGGAGAGATGATCAATTCCTTCACAACTG 120
|||
Db 227 AAACAAACAGAGAGACTGTGTGTGCTGCGGAGAGATGATCAATTCCTTCACAACTG 286
|||
QY 121 CTGC 124
|||
Db 287 CTGC 290

RESULT 11
AF164679 836 bp mRNA linear PRI 14-OCT-1999
LOCUS AF164679
DEFINITION Homo sapiens ring finger protein CKBP1 mRNA, complete cds.
ACCESSION AF164679
VERSION AF164679.1 GI:5917673
KEYWORDS
ORGANISM Homo sapiens (human)
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 836)
Son,M.Y., Park,J.-W., Kim,Y.-S., Kang,S.-W., Marshak,D.R., Park,W. and
Bae,Y.-S.
protein kinase CKII interacts with and phosphorylates the SAG
protein containing ring-H2 finger motif

JOURNAL Biochem. Biophys. Res. Commun. 263 (3), 743-748 (1999)
MEDLINE 99443734
PUBMED 10512750
REFERENCE 2 (bases 1 to 836)
AUTHORS Son,M.-Y., Park,J.-W., Kim,Y.-S., Kang,S.-W., Marshak,D.R., Park,W. and Bae,Y.-S.
TITLE Direct Submision
JOURNAL Submitted (01-JUL-1999) Department of Biochemistry, Kyungpook National University, Taegu 702-701, Korea
location/Qualifiers

FEATURES

SOURCE

1..836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
54..395
/note="SAG: contains an H2 ring finger motif"
/codon_start=1
/product="ring finger protein CKBP1"
/protein_id="AAD55984.1"
/db_xref="GI:5917674"
/translation="MADVEDGEETCALASHSSGSGSGDMPFLKKNNAVAMMSWD
VECDTCAICRYVMDACLRCAENKQEDCVVWGECHNSFHNCMSLWVKNRCPIC
QDDWVVRIGK"

CDS

BASE COUNT

228 a 185 c 223 g 200 t

ORIGIN

Query Match 100.0%; Score 124; DB 9; Length 836;
Best Local Similarity 100.0%; Pred. No. 6.5e-66;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATACGTCGCCATCTGCAGGCTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 60
|||
Db 194 CGATACGTCGCCATCTGCAGGCTCCAGGTGATGATGCTGCTTAGATGCAAGCTGA 253
|||
QY 61 AAACAAACAGAGAGACTGTGTGTGCTGCGGAGAGATGATCAATTCCTTCACAACTG 120
|||
Db 254 AAACAAACAGAGAGACTGTGTGTGCTGCGGAGAGATGATCAATTCCTTCACAACTG 313
|||
QY 121 CTGC 124
|||
Db 314 CTGC 317

RESULT 12
BD096988 754 bp DNA linear PAT 27-AUG-2002
LOCUS BD096988
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096988
VERSION BD096988.1 GI:22642576
KEYWORDS JP 2001526063-A/23.
SOURCE unidentified
ORGANISM unidentified
unclassified.
1 (bases 1 to 754)
REFERENCE Sun,Y.
SAG:apoptosis sensitivity gene
AUTHORS Patent: JP 2001526063-A 23 18-DEC-2001;
JOURNAL WARNER LAMBERT CO
OS Unidentified
PN JP 2001526063-A/23
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SRP-1998 US 60/099840.PI
YT SUN

PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
C12N5/10,C12O1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers

us-09-509-779-3_copy_141_264.rge

Db 149 GGGCCATCTGCAAGSSTCCAGGTGATGATGCGTGTCTTAATGTCGACGTGAAAAACAAC 208

Qy 69 AAGAGACTGTGTTGTGCTGGGGAGAAATGATTCATTCCTTCCACAACTGCTGC 124

RESULT 14
BD096987

LOCUS	BD096987	754 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	SAG:apoptosis sensitivity gene.				
ACCESSION	BD096987				
VERSION	BD096987.1	GI:22642575			

SOURCE ORGANISM	unidentified unidentified unclassified.
unclassified.	

AUTHORS
Sun, Y.
TITLE
SAG:apoptosis sensitivity gene
Patent: JP 2001526063-A 22 18-DEC-2001;
Matsuda T, Imamura C

COMMENT
WALKER LAWRENCE
OS Unidentified
PN JP 2001526063-A/22
PD 18-DEC-2001

PR	19-DEC-1997	US	60/068179, 11-SEP-1998	US	60/099840	PI
Y1	SUN					
PC	C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P17/02, A61P35/00,					

PC A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC
C12N5/10, C12P1/68, C12N15/00, A61K37/02, C12N5/00 CC
PC G01N33/50, G01N33/68, C12N15/00, A61K37/02, C12N5/00 CC

CC	Topology: Linear;
CC	SAG:apoptosis sensitivity gene
FH	Key Location/Qualifiers
FT	1 228
CDS	

```

FEATURES      .  FI
source        mac_peptide      1.:.339.
              _location/Qualifiers
              1.:.754
              /organism="unidentified"

```

```
BASE COUNT      206 a      153 c      202 g      193 t
ORIGIN
/mpi_type="genomic DNA"
/idx_xref="taxon:32644"
```

Query Match	90.3%	Score 112;	DB 6;	Length 754;
Best Local Similarity	100.0%	Pred. No. 2.1e-58;		
Matches 112;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 GATACGTGCGCATCTGCAGGTCAGAGTGTGATGCGCTTATGATGTCAAGCTGA 60
Db 141 CGATACGTGCGCATCTGCAGGTCAGAGTGTGATGCGCTTATGATGTCAAGCTGA 200

Qy	61	AAACAACAAGAGCATGTTGTGTGGCTGGGAGATGTAATCATTCCTTC	112
Db	201	AAACAACAAGAGCATGTTGTGTGGCTGGGAGATGTAATCATTCCTTC	252

RESULT 15	
BD096982	
LOCUS	
BD096982	
754 bp	
DNA	
1:near	
DATE 27-APR-2002	

DEFINITION
sac:apoptosis sensitively gene.
ACCESSION
BD096982
VERSION
BD096982.1 GI:22642570
KEYWORDS
JP 2001526063-A/17.

ORGANISM
unidentified
unidentified
unclassified.
1 (bases 1 to 754)

TITLE SAG:apoptosis sensitivity gene .

JOURNAL Patent: JP 2001526063-A 17 18-DEC-2001;
 COMMENT WARNER LAMBERT CO
 OS Unidentified
 PN JP 2001526063-A/17
 PD 18-DEC-2001
 PR 15-DEC-1998 JP 2000525451
 PR 19-DEC-1997 US 60/068179, 11-SEP-1998 US 60/099840 PI
 YI SUN
 PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P17/02, A61P35/00,
 PC A61P39/06, A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC
 , C12N5/10, C12O1/68,
 PC G01N33/50, G01N33/68, C12N15/00, A61K37/02, C12N5/00 CC
 Strandedness: Double;
 CC Topology: Linear;
 CC SAG:apoptosis sensitivity gene
 FH Key Location/Qualifiers
 FT CDS 1..339
 FT mat_peptide 1..339
 Location/Qualifiers
 1..754
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t

ORIGIN

Query Match 86.3%; Score 107; DB 6; Length 754;
 Best Local Similarity 100.0%; Pred. No. 2.8e-55;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GCAGGGTCAGATGATGATGCTGCTTATGATGTAAGCTGAAAACAAGAGAGACT 77
 DB 158 GCAGGGTCAGATGATGATGCTGCTTATGATGTAAGCTGAAAACAAGAGAGACT 217

QY 78 GTGTTGTGCTGGGAGAAATGTAATCATCTCTCCACAACCTGCTGC 124
 DB 218 GTGTTGTGCTGGGAGAAATGTAATCATCTCTCCACAACCTGCTGC 264

RESULT 16
 LOCUS BD096983 754 bp DNA linear PAT 27-AUG-2002
 DEFINITION SAG:apoptosis sensitivity gene.
 ACCESSION BD096983
 VERSION BD096983.1 GI:22642571
 KEYWORDS JP 2001526063-A/18.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.
 1 (bases 1 to 754)
 Sun, Y.
 SAG:apoptosis sensitivity gene
 Patent: JP 2001526063-A 18 18-DEC-2001;
 WARNER LAMBERT CO
 OS Unidentified
 PN JP 2001526063-A/18
 PD 18-DEC-2001
 PR 15-DEC-1998 JP 2000525451
 PR 19-DEC-1997 US 60/068179, 11-SEP-1998 US 60/099840 PI
 YI SUN
 PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P17/02, A61P35/00,
 PC A61P39/06, A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC
 , C12N5/10, C12O1/68,
 PC G01N33/50, G01N33/68, C12N15/00, A61K37/02, C12N5/00 CC
 Strandedness: Double;
 CC Topology: Linear;
 CC SAG:apoptosis sensitivity gene
 FH Key Location/Qualifiers
 FT CDS 1..339
 FT mat_peptide 1..339
 Location/Qualifiers

FEATURES

source 1..754
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 207 a 155 c 201 g 191 t

ORIGIN

Query Match 86.3%; Score 107; DB 6; Length 754;
 Best Local Similarity 100.0%; Pred. No. 2.8e-55;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GCAGGGTCAGATGATGATGCTGCTTATGATGTAAGCTGAAAACAAGAGAGACT 77
 DB 158 GCAGGGTCAGATGATGATGCTGCTTATGATGTAAGCTGAAAACAAGAGAGACT 217

QY 78 GTGTTGTGCTGGGAGAAATGTAATCATCTCTCCACAACCTGCTGC 124
 DB 218 GTGTTGTGCTGGGAGAAATGTAATCATCTCTCCACAACCTGCTGC 264

RESULT 17
 LOCUS BD096986 754 bp DNA linear PAT 27-AUG-2002
 DEFINITION SAG:apoptosis sensitivity gene.
 ACCESSION BD096986
 VERSION BD096986.1 GI:22642574
 KEYWORDS JP 2001526063-A/21.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.
 1 (bases 1 to 754)
 Sun, Y.
 SAG:apoptosis sensitivity gene
 Patent: JP 2001526063-A 21 18-DEC-2001;
 WARNER LAMBERT CO
 OS Unidentified
 PN JP 2001526063-A/21
 PD 18-DEC-2001
 PR 15-DEC-1998 JP 2000525451
 PR 19-DEC-1997 US 60/068179, 11-SEP-1998 US 60/099840 PI
 YI SUN
 PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P17/02, A61P35/00,
 PC A61P39/06, A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC
 , C12N5/10, C12O1/68,
 PC G01N33/50, G01N33/68, C12N15/00, A61K37/02, C12N5/00 CC
 Strandedness: Double;
 CC Topology: Linear;
 CC SAG:apoptosis sensitivity gene
 FH Key Location/Qualifiers
 FT CDS 1..339
 FT mat_peptide 1..339
 Location/Qualifiers

FEATURES

source 1..754
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 207 a 154 c 201 g 192 t

ORIGIN

Query Match 83.1%; Score 103; DB 6; Length 754;
 Best Local Similarity 100.0%; Pred. No. 9e-53;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATCGGCGGCATCTGCAGGGTCAGATGATGATGCTGCTTATGATGTAAGCTGA 60
 DB 141 CGATCGGCGGCATCTGCAGGGTCAGATGATGATGCTGCTTATGATGTAAGCTGA 200

QY 61 AAACAACAAGAGAGACTGTGTTGTGCTGGGAGAAATGTAAT 103
 DB 201 AAACAACAAGAGAGACTGTGTTGTGCTGGGAGAAATGTAAT 243

FEATURES

RESULT 18
BD096990 754 bp DNA linear PAT 27-AUG-2002
LOCUS SAG:apoptosis sensitivity gene.
DEFINITION BD096990
ACCESSION BD096990.1 GI:22642578
VERSION JP 2001526063-A/25.
KEYWORDS JP 2001526063-A/25.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 25 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/25
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FT Location/Qualifiers
FEATURES
source 1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 208 a 152 c 202 g 192 t
ORIGIN
Query Match 83.1%; Score 103; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 9e-53;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGATACGTGGCCATCTGCAGGGTCCAGGATGATGCGCTGTCTTAGATGTAACCTGA 60
DB 141 CGATACGTGGCCATCTGCAGGGTCCAGGATGATGCGCTGTCTTAGATGTAACCTGA 200
QY 61 AAACAACAAGAGACTGTGTGCTGCTGGAGAGATGTAAT 103
DB 201 AAACAACAAGAGACTGTGTGCTGCTGGAGAGATGTAAT 243
RESULT 19
BD096985 754 bp DNA linear PAT 27-AUG-2002
LOCUS SAG:apoptosis sensitivity gene.
DEFINITION BD096985
ACCESSION BD096985.1 GI:22642573
VERSION JP 2001526063-A/20.
KEYWORDS JP 2001526063-A/20.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 754)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 20 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/20
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI

YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FT Location/Qualifiers
FEATURES
source 1..754
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 206 a 155 c 201 g 192 t
ORIGIN
Query Match 78.2%; Score 97; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 5.1e-49;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGATACGTGGCCATCTGCAGGGTCCAGGATGATGCGCTGTCTTAGATGTAACCTGA 60
DB 141 CGATACGTGGCCATCTGCAGGGTCCAGGATGATGCGCTGTCTTAGATGTAACCTGA 200
QY 61 AAACAACAAGAGACTGTGTGCTGCTGGAGAGAA 97
DB 201 AAACAACAAGAGACTGTGTGCTGCTGGAGAGAA 237
RESULT 20
BD096973 747 bp DNA linear PAT 27-AUG-2002
LOCUS SAG:apoptosis sensitivity gene.
DEFINITION BD096973
ACCESSION BD096973.1 GI:22642561
VERSION JP 2001526063-A/8.
KEYWORDS JP 2001526063-A/8.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 747)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 8 18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified
PN JP 2001526063-A/8
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12Q1/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..270
FT mat_peptide 1..270.
FT Location/Qualifiers
FEATURES
source 1..747
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 204 a 155 c 197 g 191 t
ORIGIN

Query Match 72.6%; Score 90; DB 6; Length 747;
Best Local Similarity 100.0%; Pred. No. 1.2e-44;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GATGCTGCTTAGATGTCAGCTGAAACAAACAGAGAGACTGTGTGCTGGGGA 94
DB 168 GATGCTGCTTAGATGTCAGCTGAAACAAACAGAGAGACTGTGTGCTGGGGA 227
QY 95 GAATGTAATCATCTTCTCCACAACTGCTGC 124
DB 228 GAATGTAATCATCTTCTCCACAACTGCTGC 257

RESULT 21
AFJ12226 674 bp mRNA linear PRI 06-SEP-2001
LOCUS Homo sapiens SAG splice variant mRNA, complete cds.
DEFINITION AFJ12226
ACCESSION AFJ12226
VERSION AFJ12226.1 GI:13649605
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 674)
AUTHORS Duan,H., Wang,Y., Avilram,M., Swaroop,M., Loo,J.A., Bian,J.,
Tian,Y., Mueller,T., Bieganski,C.L. and Sun,Y.
SAG, a novel zinc RING finger protein that protects cells from
apoptosis induced by redox agents
Mol. Cell. Biol. 19 (4), 3145-3155 (1999)
JOURNAL MEDLINE 99182502
PUBMED 10082581
TITLE 2 (bases 1 to 674)
AUTHORS Swaroop,M., Gosink,M. and Sun,Y.
SAG/ROC2/Rbx2/Hrt2, a component of SCF E3 ubiquitin ligase: genomic
structure, a splicing variant, and two family pseudogenes
DNA Cell Biol. 20 (7), 425-434 (2001)
JOURNAL MEDLINE 21398045
PUBMED 11506706
REFERENCE 3 (bases 1 to 674)
AUTHORS Sun,Y.
Direct Submission
Submitted (09-OCT-2000) Molecular Biology, Pfizer Global Research
and Development, 2800 Plymouth Road, Ann Arbor, MI 48105, USA
JOURNAL
FEATURES
source location/Qualifiers
1..674
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3q22-q24"
1..327
/note="SAG-V; lacks zinc RING finger domain; has no
activity as a component of SCF E3 ubiquitin ligase"
/codon_start=1
/product="SAG splice variant"
/protein_id="AAK37450.1"
/db_xref="GI:13649606"
/translation="MADVDEGETCALASHSGSSGSKSGDKMFSLIKKMAVAMWSW
VECDICAIQVQVDEGIGVRNWSFALNINLASEMGFDRSGSTALAVPSVLSHSP
CLDDHR"
BASE COUNT 163 a 154 c 194 g 163 t
ORIGIN

Query Match 71.8%; Score 89; DB 9; Length 674;
Best Local Similarity 100.0%; Pred. No. 5.2e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 ATGCTGCTTAGATGTCAGCTGAAACAAACAGAGAGACTGTGTGCTGGGGA 95
DB 440 ATGCTGCTTAGATGTCAGCTGAAACAAACAGAGAGACTGTGTGCTGGGGA 499
QY 96 AATGTAATCATCTTCTCCACAACTGCTGC 124

DB 500 AATGTAATCATCTTCTCCACAACTGCTGC 528

RESULT 22
BD096984 754 bp DNA linear PAT 27-AUG-2002
LOCUS BD096984
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096984
VERSION BD096984.1 GI:22642572
KEYWORDS JP 2001526063-A/19.
SOURCE unidentified
ORGANISM unidentified
REFERENCE unclassified.
1 (bases 1 to 754)
AUTHORS Sun,Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 19-18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified

QY 42 GTCTTAGATGTCAGCTGAAACAAACAGAGAGACTGTGTGCTGGGGAATGTA 101
DB 182 GTCTTAGATGTCAGCTGAAACAAACAGAGAGACTGTGTGCTGGGGAATGTA 241
QY 102 ATCATCTCTTCCACAACTGCTGC 124
DB 242 ATCATCTCTTCCACAACTGCTGC 264

RESULT 23
BD096992 754 bp DNA linear PAT 27-AUG-2002
LOCUS BD096992
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096992
VERSION BD096992.1 GI:22642580
KEYWORDS JP 2001526063-A/27.
SOURCE unidentified
ORGANISM unidentified
REFERENCE unclassified.
1 (bases 1 to 754)
AUTHORS Sun,Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 27-18-DEC-2001;
WARNER LAMBERT CO
COMMENT OS Unidentified

Query Match 66.9%; Score 83; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 3e-40;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BASE COUNT 206 a 155 c 201 g 192 t
ORIGIN

FEATURES
source location/Qualifiers
1..754
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

PN JP 2001526063-A/27
PD 18-DEC-2001
PF 15-DEC-1998 JP 200525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C1201/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FEATURES
source
1..754
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 206 a 155 c 201 g 192 t
ORIGIN
Query Match 61.3%; Score 76; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 7,2e-36;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATACGTGCGCCATCTGCAGGCTCCAGGTGATGATGCTGCTTATGATGCAAGCTCA 60
141 CCATACGTGCGCCATCTGCAGGCTCCAGGTGATGATGCTGCTTATGATGCAAGCTCA 200
DB 61 AAACAACAAGAGGAC 76
201 AAACAACAAGAGGAC 216
QY
DB
RESULT 24
BD096991 754 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096991.1 GI:22642579
VERSION BD096991.1
KEYWORDS JP 2001526063-A/26.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 754)
Sun,Y.
SAG:apoptosis sensitivity gene
Patent: JP 2001526063-A 26 18-DEC-2001;
WARNER LAMBERT CO
OS Unidentified
PN JP 2001526063-A/26
PD 18-DEC-2001
PF 15-DEC-1998 JP 200525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C1201/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FEATURES
source
1..754
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 206 a 155 c 201 g 192 t
ORIGIN
Query Match 59.7%; Score 74; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 1,3e-34;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 GTCAAGCTGAACAAACAAGAGGACTGTGTTGCTGCTGGGAGAAATGATCATTCCT 110
191 GTCAAGCTGAACAAACAAGAGGACTGTGTTGCTGCTGGGAGAAATGATCATTCCT 250
DB 111 TCCACAACCTGCTGC 124
251 TCCACAACCTGCTGC 264
QY
DB
RESULT 25
BD096993 754 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096993.1 GI:22642581
VERSION BD096993.1
KEYWORDS JP 2001526063-A/28.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 754)
Sun,Y.
SAG:apoptosis sensitivity gene
Patent: JP 2001526063-A 28 18-DEC-2001;
WARNER LAMBERT CO
OS Unidentified
PN JP 2001526063-A/28
PD 18-DEC-2001
PF 15-DEC-1998 JP 200525451
PR 19-DEC-1997 US 60/068179,11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C1201/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC SAG:apoptosis sensitivity gene
FH Key Location/Qualifiers
FT CDS 1..339
FT mat_peptide 1..339.
FEATURES
source
1..754
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 207 a 155 c 201 g 191 t
ORIGIN
Query Match 59.7%; Score 74; DB 6; Length 754;
Best Local Similarity 100.0%; Pred. No. 1,3e-34;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 GTCAAGCTGAACAAACAAGAGGACTGTGTTGCTGCTGGGAGAAATGATCATTCCT 110
191 GTCAAGCTGAACAAACAAGAGGACTGTGTTGCTGCTGGGAGAAATGATCATTCCT 250
DB 111 TCCACAACCTGCTGC 124
251 TCCACAACCTGCTGC 264
QY
DB
RESULT 26
AC112771 35638 bp DNA linear PRI 28-SEP-2002
LOCUS

DEFINITION	Homo sapiens 3 RAC RP11-144C9 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION	AC112771
VERSION	AC112771.6
KEYWORDS	GI:22003982
SOURCE	HTG.
ORGANISM	Homo sapiens (human)
REFERENCE	Mammalia; Euteleia; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 35638) Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralungue,H.C., Are,J.R., Ayale,M., Banks,T., Barbakaj,J., Benton,J., Bimaga,K., Blankenburg,K., Bonini,D., Bowck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Donahewite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,D., Garcia,A., Garner,T., Garza,N., Gill,R., Gorelli,J.H., Guvera,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulvik,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korva,J., Kovari,C., Katovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogun,M., Okmonou,G., Oragnye,N., Oviedo,R., Pace,A., Payton,B., Peery,D., Perez,L., Peters,L., Plickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rotubokan,I., Rolfe,M., Ruiz,S., Savory,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,B., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Stacke,A., Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Teitrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and Gibbs,R.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 35638)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (24-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 35638)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (25-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4 (bases 1 to 35638)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	5 (bases 1 to 35638)
AUTHORS	Worley,K.C.

TITLE	Direct Submission
JOURNAL	Submitted (28-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Jul 30, 2002 this sequence version replaced gi:20976460. INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

FEATURES	SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.
SOURCE	QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html .
misc_feature	Location/Qualifiers 1..35638 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="3" /clone="RP11-144C9" 1..1895 /note="overlaps bases 56107..58001 of clone AC092978"
STS	/function="Clone overlap" 14..341 /standard_name="SHGC-105132" complement(1283..1580) /rpt_family="AluSkx" 1618..1906 /rpt_family="AluSkx" 1907..1986 /rpt_family="(GGGGA)n" complement(2109..2203) /rpt_family="AluM4" complement(2204..2463) /rpt_family="AluSkx" complement(2464..2700) /rpt_family="AluM4" complement(2701..2819) /rpt_family="AluSc" complement(2843..3559) /rpt_family="AluM4" complement(3541..3642) /rpt_family="AluSg/x" 3643..3758 repeat_region

```

repeat_region      /rpt_family="FLAM_A"
                    complement(3777..4993)
                    /rpt_family="LIMEC"
repeat_region      complement(4994..5287)
                    /rpt_family="Alusg"
repeat_region      complement(5288..5404)
                    /rpt_family="LIMEC"
repeat_region      complement(5405..5716)
                    /rpt_family="Alusx"
repeat_region      complement(5717..6153)
                    /rpt_family="LIMEC"
repeat_region      complement(6154..6379)
                    /rpt_family="Alusg/x"
repeat_region      complement(6380..6394)
                    /rpt_family="LIMEC"
repeat_region      complement(6427..6570)
                    /rpt_family="MER91A"
repeat_region      7418..7561
                    /rpt_family="Alusg/x"
repeat_region      8094..8139
                    /rpt_family="(TG)n"
repeat_region      complement(10186..10802)
                    /rpt_family="L2"
repeat_region      11949..12489
                    /rpt_family="HAL1"
repeat_region      12490..12772
                    /rpt_family="Alusx"
repeat_region      12773..12804
                    /rpt_family="(TAAA)n"
repeat_region      12805..12820
                    /rpt_family="Alusx"
repeat_region      12821..12896
                    /rpt_family="HAL1"
repeat_region      12989..13063
                    /rpt_family="Alusg/x"
repeat_region      13069..13452
                    /rpt_family="HAL1"
repeat_region      13453..13585
                    /rpt_family="FLAM_C"
repeat_region      13586..13609
                    /rpt_family="HAL1"
repeat_region      13619..13654
                    /rpt_family="G-rich"
repeat_region      13653..13809
                    /rpt_family="GA-rich"
repeat_region      complement(13819..14048)
                    /rpt_family="AluJo"
repeat_region      14077..14170
                    /rpt_family="MER113"
repeat_region      complement(14655..14830)
                    /rpt_family="MUT1J"
repeat_region      complement(14852..15139)
                    /rpt_family="Alusp"

```

Query Match 39.5%; Score 49; DB 9; Length 35638;
 Best Local Similarity 100.0%; Pred. No. 7.4e-19;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GATGCTGTTAGATGTCAGCTGAAACAAACAGAGACTGTGTTG 83
 Db 30671 GATGCTGTTAGATGTCAGCTGAAACAAACAGAGACTGTGTTG 30719

RESULT 27
 AC068693/c 125041 bp DNA 1linear HTG 04-MAY-2001
 DEFINITION Homo sapiens chromosome 17 clone RP11-144C9 map 17, WORKING DRAFT
 AC068693
 ACCESSION AC068693.2 GI:8139716
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS Homo sapiens (human)
 SOURCE ORGANISM Homo sapiens

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 125041)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 17, clone RP11-144C9
 Unpublished
 2 (bases 1 to 125041)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Baeten, V., Bedalov, F.,
 Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
 Campopiano, A., Castle, A., Choepiel, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heatord, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lechoczky, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
 Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisanil, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testafaye, S., Theodore, J., Tirrell, A., Travers, M., Triggilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 1, 2000 this sequence version replaced gi:7712215.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: 144 C 9
 Center clone name: 144 C 9
 Summary Statistics
 Sequencing vector: M13; W77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 112098 bases at least Q40
 Consensus quality: 118849 bases at least Q40
 Consensus quality: 121119 bases at least Q20
 Insert size: 158000; agarose-fp
 Insert size: 122641; sum-of-contigs
 Quality coverage: 3.8 in Q20 bases; agarose-fp
 Quality coverage: 4.9 in Q20 bases; sum-of-contigs

 NOTE: This is a 'working draft' sequence. It currently
 consists of 25 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 1019: contig of 1019 bp in length
 * 1020 1119: gap of 100 bp
 * 1120 3013: contig of 1894 bp in length
 * 3014 3113: gap of 100 bp
 * 3114 5008: contig of 1895 bp in length
 * 5009 5108: gap of 100 bp
 * 5109 7154: contig of 2046 bp in length
 * 7155 7254: gap of 100 bp
 * 7255 10405: contig of 3151 bp in length
 * 10406 10505: gap of 100 bp

* 10506 12914: contig of 2409 bp in length
* 12915 13014: gap of 100 bp
* 13015 13781: contig of 767 bp in length
* 13782 13881: gap of 100 bp
* 13882 17746: contig of 3865 bp in length
* 17747 17846: gap of 100 bp
* 17847 20494: contig of 2648 bp in length
* 20495 20594: gap of 100 bp
* 20595 22173: contig of 1579 bp in length
* 22174 22273: gap of 100 bp
* 22274 25788: contig of 3515 bp in length
* 25789 30218: gap of 100 bp
* 30219 30318: contig of 4330 bp in length
* 30319 33343: gap of 100 bp
* 33344 33443: gap of 100 bp
* 33444 36428: contig of 2985 bp in length
* 36429 36528: gap of 100 bp
* 36529 40028: contig of 3500 bp in length
* 40029 40128: gap of 100 bp
* 40129 45543: contig of 5415 bp in length
* 45544 45643: gap of 100 bp
* 45644 52370: contig of 6727 bp in length
* 52371 52470: gap of 100 bp
* 52471 56869: contig of 4399 bp in length
* 56870 62452: gap of 100 bp
* 62453 62552: gap of 100 bp
* 62553 70768: contig of 8216 bp in length
* 70769 70868: gap of 100 bp
* 70869 76175: contig of 5307 bp in length
* 76176 76275: gap of 100 bp
* 76276 84628: contig of 8353 bp in length
* 84629 84729: gap of 100 bp
* 84729 95400: contig of 10672 bp in length
* 95401 95500: gap of 100 bp
* 95501 111561: contig of 16061 bp in length
* 111562 111662: gap of 100 bp
* 111662 125041: contig of 13380 bp in length.

FEATURES

source

1. .125041 /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-144C9"
/clone_lib="RPC1-11 Human Male BAC"
1. .1019 /note="assembly_fragment"
1120. .3013 /note="assembly_fragment"
3114. .5008 /note="assembly_fragment"
5109. .7154 /note="assembly_fragment"
7255. .10405 /note="assembly_fragment"
10506. .12914 /note="assembly_fragment"
13015. .13781 /note="assembly_fragment"
13882. .17746 /note="assembly_fragment"
17847. .20494 /note="assembly_fragment"
20595. .22173 /note="assembly_fragment"
22174. .25788 /note="assembly_fragment"
25789. .30218 /note="assembly_fragment"
30219. .30318 /note="assembly_fragment"
30319. .33343 /note="assembly_fragment"
33344. .33443 /note="assembly_fragment"
33444. .36428 /note="assembly_fragment"
36429. .36528 /note="assembly_fragment"
36529. .40028 /note="assembly_fragment"
40029. .40128 /note="assembly_fragment"
40129. .45543 /note="assembly_fragment"
45544. .45643 /note="assembly_fragment"
45644. .52370 /note="assembly_fragment"
52371. .52470 /note="assembly_fragment"
52471. .56869 /note="assembly_fragment"
56870. .62452 /note="assembly_fragment"
62453. .62552 /note="assembly_fragment"
62553. .70768 /note="assembly_fragment"
70769. .70868 /note="assembly_fragment"
70869. .76175 /note="assembly_fragment"
76176. .76275 /note="assembly_fragment"
76276. .84628 /note="assembly_fragment"
84629. .84729 /note="assembly_fragment"
84729. .95400 /note="assembly_fragment"
95401. .95500 /note="assembly_fragment"
95501. .111561 /note="assembly_fragment"
111562. .111662 /note="assembly_fragment"
111662. .125041 /note="assembly_fragment"

misc_feature /note="assembly_fragment"
25889. .30218 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
30319. .33343 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
33444. .36428 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
36529. .40028 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
40129. .45543 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
45644. .52370 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
52471. .56869 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
56970. .62452 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
62553. .70768 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
70869. .76175 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
76276. .84628 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
84729. .95400 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
95501. .111561 /note="assembly_fragment"
misc_feature /note="assembly_fragment"
111662. .125041 /note="assembly_fragment"

BASE COUNT 34737 a 26838 c 26914 g 34145 t 2407 others
ORIGIN

Query Match 39.5%; Score 49; DB 2; Length 125041;
Best Local Similarity 100.0%; Pred. No. 8 1e-19;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GATCCCTGCTTAAATGTCAGCTGAACCAACAGAGACTGTGTTG 83

Db 17697 GATCCCTGCTTAAATGTCAGCTGAACCAACAGAGACTGTGTTG 17609

RESULT 28
LOCUS AC106176/c 238330 bp DNA linear HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-20P16, WORKING DRAFT SEQUENCE, 3
unordered pieces.
ACCESSION AC106176 GI:30578781
VERSION AC106176.5
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 238330)
AUTHORS Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, B., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, B.,
Cardenas, V., Carter, K., Cavazos, I., Cessat, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Girdy, M., Guerra, M., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

TITLE	JOURNAL	REFERENCE	AUTHORS
Direct Submission Unpublished 2 (bases 1 to 238330)			Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huliyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacod, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowalski, C., Kraft, C. L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louiseged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinec, E., Mathewey, S., McLeod, M. P., McNelly, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokilemb, O., Okunnu, G., Olarpunsaagoo, A., Pal, S., Parke, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pflankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quitoz, J., Rachlin, E., Reeves, K., Reiger, M. A., Relg, R., Rellly, B., Rellly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Severy, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Sheely, J., Shvartsbeyn, A., Slisson, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trijos, Z., Umani, K., Valias, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlecyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederstock, A., Weiser, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.
Direct Submission Submitted (12-JAN-2002) Of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 3 (bases 1 to 238330) Rat Genome Sequencing Consortium. Direct Submission Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One May 13, 2003 This sequence version replaced gi:24819215. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly ('contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.			
COMMENT			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu Project Information Center project name: GJOH Center clone name: CH230-20P16 Summary Statistics Assembly program: Atlas 3.0; Consensus quality: 229461 bases at least Q40 Consensus quality: 231714 bases at least Q30 Consensus quality: 233148 bases at least Q20 Estimated insert size: 239075; sum-of-contigs estimation Quality coverage: 8x in Q20 bases; sum-of-contigs estimation			

FEATURES	source
* NOTE: Estimated insert size may differ from sequence length	
* See http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html .	
* NOTE: This is a 'working draft' sequence. It currently	
* consists of 3 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
* 1 236056: contig of 236056 bp in length	
* 236057 236156: gap of unknown length	
* 236157 237213: contig of 1057 bp in length	
* 237214 237313: gap of unknown length	
* 237314 238330: contig of 1017 bp in length.	
* Location/Qualifiers	
1. 238330	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10116"	
/clone="CH230-20P16"	
1. 1582	
/note="wgs_contig"	
BASE COUNT 64585 a 53668 c 53820 g 62326 t 3931 others	
ORIGIN	
Query Match 39.5%; Score 49; DB 2; Length 238330;	
Best Local Similarity 100.0%; Pred. No. 8,5e-19;	
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
35 GATCCGCTGCTGTGATGTCAAGCTCAAGCAAAACCAAGAGCAGCTGTG 83	
Db 158930 GATCCCTGCTGTGATGTCAAGCTCAAGCAAAACCAAGAGCAGCTGTG 158882	
RESULT 29	
AC098496/c	
LOCUS	
DEFINITION	
AC098496 245476 bp DNA linear HTG 10-MAY-2003	
Rattus norvegicus clone CH230-23C21, WORKING DRAFT SEQUENCE, 4	
unordered pieces.	
AC098496	
AC098496.8 GI:30521845	
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.	
Rattus norvegicus (Norway rat)	
Rattus norvegicus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
Rattus.	
1 (bases 1 to 245476)	
Mizny,D,Marle,, Metzker,M,Lee,, Abramson,S,, Adams,C,, Alder,J,,	
Allen,C,, Allen,H,, Aishrooks,S,, Amin,A,, Anguiano,D,,	
AyalaDebechi,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,,	
Baldwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F,,	
Biswalo,K,, Blair,J,, Blankenburg,K,, Blyth,P,, Brown,M,,	
Bryan,N,, Buhey,C,, Burch,P,, Burrell,K,, Calderon,E,,	
Cardenas,V,, Carter,K,, Cavazos,I,, Ceasar,H,, Center,A,,	
Chacho,J,, Chavez,D,, Chen,G,, Chen,R,, Chen,Y,, Chen,Z,, Chu,J,,	
Cleveland,C,, Cockrell,R,, Cox,C,, Coyle,M,, Cree,A,, D'Souza,L,,	
Devilla,M,L,, Davis,C,, Davy-Carroll,L,, De Anda,C,, Dederich,D,,	
Delgado,O,, Denson,S,, Derram,C,, Ding,Y,, Dinh,H,, Diya,K,,	
Draper,H,, Dugan-Rocha,S,, Dunn,A,, Dublin,K,, Duval,B,, Eaves,K,,	
Egan,A,, Escotto,M,, Eugene,C,, Evans,C,A,, Falle,T,, Fan,G,,	
Fernandez,S,, Finley,M,, Flagg,N,, Forbes,L,, Foster,M,, Foster,P,,	
Fraser,C,M,, Gabris,A,, Ganta,R,, Garcia,A,, Garner,T,, Garza,M,,	
Gebregorzi,E,, Geer,K,, Gill,R,, Grady,M,, Guerra,M,, Guevara,W,,	
Gunararatne,P,, Haaland,W,, Hamil,C,, Hamilton,N,, Hamilton,K,,	
Harvey,Y,, Havlak,P,, Hawes,A,, Henderson,N,, Hernandez,J,,	
Hernandez,R,, Hines,S,, Hladun,S,L,, Hodgson,A,, Hognes,M,,	
Hollins,B,, Howell,S,, Huylk,S,, Hume,J,, Idlebird,A,, Jackson,A,,	
Jackson,L,, Jacob,L,, Jiang,H,, Johnson,B,, Johnson,R,, Jolivet,A,,	
Karpachy,S,, Kelly,S,, Kelly,S,, Khan,Z,, King,L,, Koyar,C,,	
Kowitz,C,, Kraft,C,L,, Lebow,H,, Levan,J,, Lewis,L,, Li,Z,, Liu,J,,	
REFERENCE	
AUTHORS	

Li, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensen, L., Louised, H., Lozano, R., Lu, X., Ma, J.,
 Maheshwari, W., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Manthey, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Narkervis, C., Neal, D., Newton, G., Nguyen, N., Norris, S.,
 Nwankemeh, O., Okonmu, G., Olarunpungon, A., Pal, S., Parke, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smaj, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Vals, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczek, R., Woodson, H., Morley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, U., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G., and Gibbs, R.A.

Unpublished
 Direct Submissions
 2 (bases 1 to 245476)
 Worley, K.C.

REFERENCE
 JOURNAL
 TITLERS
 AUTHORS
 JOURNAL
 Submitted (24-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 245476)
 Rat Genome Sequencing Consortium.

REFERENCE
 JOURNAL
 TITLERS
 AUTHORS
 JOURNAL
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24819573.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: GHEY
 Center clone name: CH230-23C21

----- Summary Statistics
 Assembly program: Atlas 3.0
 Consensus quality: 228381 bases at least Q40
 Consensus quality: 230024 bases at least Q30
 Consensus quality: 231602 bases at least Q20
 Estimated insert size: 236545; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces

is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 5516: contig of 5516 bp in length
 5517 5616: gap of unknown length
 5617 242823: contig of 237207 bp in length
 242824 242923: gap of unknown length
 242924 243958: contig of 1035 bp in length
 243959 244058: gap of unknown length
 244059 245476: contig of 1418 bp in length.

Location/Qualifiers
 1. 245476
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-23C21"
 1. 1560
 /note="wgs end extension
 clone_end:T7
 complement(3889..4558)
 /note="clone boundary
 clone_end:T7
 site:EcoRI
 end_sequence:BH61779"
 5617..6729
 /note="wgs contig"
 69231..103385
 /note="clone boundary
 clone_end:Sp6
 site:EcoRI
 end_sequence:BH61780"

BASE COUNT 67659 a 50656 c 50781 g 63728 t 12652 others
 ORIGIN

Query Match 39.5%; Score 49; DB 2; Length 245476;
 Best Local Similarity 100.0%; Pred. No. 8.5e-19;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GATGCTGTCTTGTAGATGTCAGGCTGAACCAACAGAGAGCTGTGTTG 83
 DB 64017 GATGCTGTCTTGTAGATGTCAGGCTGAACCAACAGAGAGCTGTGTTG 63969

RESULT 30
 AC118192/c 187353 bp DNA linear HTG 14-APR-2002
 LOCUS
 DEFINITION Mus musculus clone RP23-200C17, WORKING DRAFT SEQUENCE, 10 ordered
 pieces.
 AC118192
 VERSION AC118192.1 GI:20147957
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULFILLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 187353)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Mus musculus, clone RP23-200C17

Unpublished
 2 (bases 1 to 187353)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barre, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Camarata, J., Campobiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A.,
 Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fardo, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gairyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamet, A., Karlas, A., Kelle, C., Lavoie, K., Lamazares, R.,
 Landers, T., Lenocky, J., Levine, R., Lindblad-Toh, K., Liu, G.,

TITLE Direct Submission
JOURNAL Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:

```

CC id W06951
CC est
CC blastn
CC identity 96
CC region 7.64
CC id W06951
CC est
CC blastn
CC identity 97
CC region 2.179
CC id W38711
CC est
CC Von Heijne matrix
CC score 3.7
CC seq ETCALASHSGSSG/SK
CC n-a, g, c or t
FH Key Location/Qualifiers
FT misc_feature 89..222
FT misc_feature 26..90
FT misc_feature 89..222
FT misc_feature 33..90
FT misc_feature 45..222
FT sig_peptide 24..86
FT misc_feature 3..
Location/Qualifiers
1..224
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

BASE COUNT 41 a 60 c 83 g 38 t 2 others
ORIGIN

```

```

Query Match 34.7% Score 43; DB 6; Length 224;
Best Local Similarity 100.0%; Pred. No. 3e-15;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CGATACGTGCGCATCTGCAGGATCCAGGTGATGATGCTGT 43
Db 164 CGATACGTGCGCATCTGCAGGATCCAGGTGATGATGCTGT 206

```

```

RESULT 32
AC104983
LOCUS
DEFINITION Homo sapiens chromosome 18 clone RP11-352C3 map 18, LOW-PASS
SEQUENCE SAMPLING.
AC104983 59155 bp DNA linear HTG 22-DEC-2001
AC104983.1 GI:17977361
HTG; HTGS_PHASEO.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 59155)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-352C3
Unpublished
2 (bases 1 to 59155)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campolano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gaidyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Headford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kelle,C., Lacroque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McSheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N., Pollara,V.,

```

TITLE JOURNAL COMMENT

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zalnoun,J., Zemdek,L., Zimmer,A. and Zody,M.

Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L22883
Center clone name: 352_C_3

* NOTE: This record contains 70 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

1 725 824: contig of 724 bp in length
825 824: gap of 100 bp
1583 1582: contig of 758 bp in length
1683 1682: gap of 100 bp
2409 2409: contig of 727 bp in length
2510 2509: gap of 100 bp
3235 3235: contig of 726 bp in length
3336 3335: gap of 100 bp
4091 4090: contig of 755 bp in length
4191 4190: gap of 100 bp
4955 4955: contig of 765 bp in length
4956 4955: gap of 100 bp
5056 5055: gap of 100 bp
5780 5780: contig of 725 bp in length
5881 5880: gap of 100 bp
6644 6643: contig of 763 bp in length
6744 6743: gap of 100 bp
7496 7496: contig of 753 bp in length
7497 7496: gap of 100 bp
7597 7596: gap of 100 bp
8349 8348: contig of 752 bp in length
8448 8448: gap of 100 bp
9217 9217: contig of 769 bp in length
9317 9317: gap of 100 bp
9318 9317: gap of 100 bp
10093 10093: contig of 776 bp in length
10094 10093: gap of 100 bp
10194 10193: gap of 100 bp
10924 10924: contig of 731 bp in length
11024 11024: gap of 100 bp
11793 11793: contig of 769 bp in length
11794 11793: gap of 100 bp
11893 11893: gap of 100 bp
11894 11894: contig of 768 bp in length
12662 12661: gap of 100 bp
12762 12761: gap of 100 bp
13526 13525: contig of 764 bp in length
13626 13625: gap of 100 bp
14348 14347: contig of 722 bp in length
14447 14447: gap of 100 bp
14488 14487: gap of 100 bp
15173 15172: contig of 725 bp in length
15173 15172: gap of 100 bp
15273 15272: gap of 100 bp
16008 16007: contig of 735 bp in length
16108 16107: gap of 100 bp
16855 16854: contig of 748 bp in length
16855 16854: gap of 100 bp

```

```

* 16956 17711: contig of 756 bp in length
* 17712 17811: gap of 100 bp
* 17812 18568: contig of 757 bp in length
* 18568 18668: gap of 100 bp
* 18668 19421: contig of 753 bp in length
* 19421 19521: gap of 100 bp
* 19521 20229: contig of 708 bp in length
* 20229 20329: gap of 100 bp
* 20329 21061: contig of 732 bp in length
* 21061 21161: gap of 100 bp
* 21161 21909: contig of 748 bp in length
* 21909 22009: gap of 100 bp
* 22009 22773: contig of 764 bp in length
* 22773 22873: gap of 100 bp
* 22873 23623: contig of 750 bp in length
* 23623 23723: gap of 100 bp
* 23723 24489: contig of 766 bp in length
* 24489 24590: gap of 100 bp
* 24590 25354: contig of 765 bp in length
* 25354 25454: gap of 100 bp
* 25454 26174: contig of 720 bp in length
* 26174 26274: gap of 100 bp
* 26274 27043: contig of 769 bp in length
* 27043 27143: gap of 100 bp
* 27143 27900: contig of 757 bp in length
* 27900 28000: gap of 100 bp
* 28000 28753: contig of 753 bp in length
* 28753 28853: gap of 100 bp
* 28853 29572: contig of 719 bp in length
* 29572 29672: gap of 100 bp
* 29672 30419: contig of 747 bp in length
* 30419 30519: gap of 100 bp
* 30519 31262: contig of 743 bp in length
* 31262 31362: gap of 100 bp
* 31362 32052: contig of 690 bp in length
* 32052 32152: gap of 100 bp
* 32152 32909: contig of 757 bp in length
* 32909 33009: gap of 100 bp
* 33009 33736: contig of 727 bp in length
* 33736 33836: gap of 100 bp
* 33836 34598: contig of 762 bp in length
* 34598 34698: gap of 100 bp
* 34698 35456: contig of 758 bp in length
* 35456 35556: gap of 100 bp
* 35556 36274: contig of 718 bp in length
* 36274 36374: gap of 100 bp
* 36374 37111: contig of 737 bp in length
* 37111 37211: gap of 100 bp
* 37211 37954: contig of 743 bp in length
* 37954 38054: gap of 100 bp
* 38054 38783: contig of 725 bp in length
* 38783 38883: gap of 100 bp
* 38883 39686: contig of 803 bp in length
* 39686 39786: gap of 100 bp
* 39786 40535: contig of 749 bp in length
* 40535 40635: gap of 100 bp
* 40635 41387: contig of 752 bp in length
* 41387 41487: gap of 100 bp
* 41487 42254: contig of 767 bp in length
* 42254 42354: gap of 100 bp
* 42354 43046: contig of 692 bp in length
* 43046 43146: gap of 100 bp
* 43146 43885: contig of 735 bp in length
* 43885 43985: gap of 100 bp
* 43985 44741: contig of 756 bp in length
* 44741 44841: gap of 100 bp
* 44841 45604: contig of 763 bp in length
* 45604 45704: gap of 100 bp
* 45704 46429: contig of 725 bp in length
* 46429 46529: gap of 100 bp
* 46529 47244: contig of 715 bp in length
* 47244 47344: gap of 100 bp
* 47344 48088: contig of 744 bp in length

```

```

* 48089 48188: gap of 100 bp
* 48188 48948: contig of 760 bp in length
* 48948 49048: gap of 100 bp
* 49048 49803: contig of 755 bp in length
* 49803 49903: gap of 100 bp
* 49903 50648: contig of 745 bp in length
* 50648 50748: gap of 100 bp
* 50748 51506: contig of 758 bp in length
* 51506 51606: gap of 100 bp
* 51606 52345: contig of 739 bp in length
* 52345 52445: gap of 100 bp
* 52445 53185: contig of 740 bp in length
* 53185 53285: gap of 100 bp
* 53285 54059: contig of 774 bp in length
* 54059 54159: gap of 100 bp
* 54159 54901: contig of 742 bp in length
* 54901 55001: gap of 100 bp
* 55001 55740: contig of 739 bp in length
* 55740 55840: gap of 100 bp
* 55840 56577: contig of 737 bp in length
* 56577 56677: gap of 100 bp
* 56677 57433: contig of 756 bp in length
* 57433 57533: gap of 100 bp
* 57533 58290: contig of 757 bp in length

Query Match
Best Local Similarity 34.7%; Score 43; DB 2; Length 59155;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 AGATGTCAGCTGAACACACAGAGAGCTGTGTGCTCT 89
Db 7777 AGATGTCAGCTGAACACACAGAGAGCTGTGTGCTCT 7819

RESULT 33
AC108040/c
LOCUS
DEFINITION
Homoe sapiens chromosome 4 clone RP11-210010, WORKING DRAFT
AC108040
AC108040.2 GI:18425316
VERSION
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULFOP.
KEYWORDS
SOURCE
Homoe sapiens (human)
ORGANISM
Homoe sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 163521)
AUTHORS
Waterston, R.H.
JOURNAL
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 163521)
AUTHORS
Waterston, R.H.
JOURNAL
Direct Submission
Submitted (24-JAN-2002) Genome Sequencing Center, Washington
University, School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jan 30, 2002 this sequence version replaced gi:18308830.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
Project Information
Center project name: H.NH0210010
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-Primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 162304 bases at least Q40

```


Consensus quality: 162444 bases at least Q30
 Consensus quality: 162529 bases at least Q20
 Insert size: 189000; agarose-fp
 Insert size: 163221; sum-of-contigs
 Quality coverage: 8.88 in Q20 bases; agarose-fp
 Quality coverage: 8.85 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1564: contig of 1564 bp in length
 * 1565 1664: gap of unknown length
 * 1665 88984: contig of 87320 bp in length
 * 88985 89084: gap of unknown length
 * 89085 112702: contig of 23618 bp in length
 * 112703 112802: gap of unknown length
 * 112803 163521: contig of 50719 bp in length.

FEATURES
 source
 1. 163521
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="RP11-210010"

misc_feature
 1. 1564
 /note="assembly_name:Contig25
 clone_end:SP6
 vector_side:left"

misc_feature
 1665. 88984
 /note="assembly_name:Contig28
 clone_end:T7
 vector_side:left"

misc_feature
 89085. 112702
 /note="assembly_name:Contig26"
 112803. 163521
 /note="assembly_name:Contig27"

BASE COUNT 51049 a 32426 c 33346 g 46399 t 301 others
 ORIGIN

Query Match 34.7%; Score 43; DB 2; Length 163521;
 Best Local Similarity 100.0%; Pred. No. 4.7e-15;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 ACATGTCACAGCTGAAACAAACAGAGAGCTGTTGTGCTCT 89
 |||||
 Db 80232 ACATGTCACAGCTGAAACAAACAGAGAGCTGTTGTGCTCT 80190

RESULT 34
 AC092418/c 164500 bp DNA linear PRI 18-DEC-2002
 LOCUS AC092418 Homo sapiens chromosome 3 clone RP11-229A12, complete sequence.
 DEFINITION AC092418 AC016923
 AC092418.3 GI:27228861
 VERSION HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
 Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
 Haugen, E.D.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 164500)
 AUTHORS Kaul, R.K., Olson, M.V., Raymond, C., Clendenning, J., Ivey, R.G. and
 Haugen, E.D.

TITLE Direct Submission
 JOURNAL Submitted (04-JUL-2001) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 REFERENCE 3 (bases 1 to 164500)
 AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
 Saenphimmachak, C., Phelps, K.A., Buckley, D., Raymond, C. and
 Haugen, E.D.

TITLE Direct Submission
 JOURNAL Submitted (24-MAY-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 REFERENCE 4 (bases 1 to 164500)
 AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
 Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
 Haugen, E.D.

TITLE Direct Submission
 JOURNAL Submitted (18-DEC-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 COMMENT On Dec 18, 2002 this sequence version replaced gi:21166206.

Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgc@gsu.washington.edu
 Drafting Center: BCM

----- Project Information
 Center project name: chr-3
 Center clone name: RP11-229A12 (bc0315)

----- Summary Statistics
 Sequencing vector: unknown; 49% of reads
 Sequencing vector: plasmid; 51% of reads
 Chemistry: Dye-terminator ET; 48% of reads
 Chemistry: Unknown; 0% of reads
 Chemistry: Dye-terminator Big Dye; 51% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 164200 bases at least Q40
 Consensus quality: 164447 bases at least Q30
 Consensus quality: 164498 bases at least Q20
 Insert size: 164500; sum-of-contigs
 Quality coverage: 7.7x in Q20 bases; sum-of-contigs

----- Overlapping Sequences:
 5': RP11-120C2 (UWGC:bc0239) AC121250, 4576-bp overlap
 3': RP11-755B10 (UWGC:bc0671) AC114480

----- Sequence Quality Assessment:
 This entry has been annotated with sequence quality
 estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than
 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the
 GenBank flat file format but are available as part
 of this entry's ASN.1 file.

----- This sequence was finished as follows unless otherwise noted:
 all regions were either double-stranded or sequenced with an
 alternate chemistry or covered by high quality data (i.e., Phred
 quality >= 30); an attempt was made to resolve all sequencing
 problems, such as compressions and repeats; all regions were
 covered by at least one plasmid subclone or more than one M13
 subclone; and the assembly was confirmed by restriction digest.

----- Sequence Validation:
 This sequence has been validated by Multiple Complete Digest
 fingerprinting. Comparison of the experimentally derived digest
 fragments with sequence-predicted fragments is given below.
 The electronically-digested sequence consists of both insert and
 vector, in order to accurately represent the entire circular BAC.
 Small fragments below a variable cutoff (approximately 400-800 bp)
 are not resolved in the fingerprint and hence do not appear
 in the table. There are no significant remaining discrepancies
 between the experimental and predicted values. Uniquely ordered
 fragments are separated by dashed lines.

EcoRI		HindIII		BglII	
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
-----	-----	-----	-----	-----	-----
8696	8826	826	813	4194	4490
-----	-----	-----	-----	-----	-----
6	<800	6382	6597	2067	2161
-----	-----	-----	-----	-----	-----
4056	4085	512	<800	7354	7474
-----	-----	-----	-----	-----	-----
2246	2301	449	<800	13554	13505
-----	-----	-----	-----	-----	-----
4995	4994	1929	1902	1393	1424
-----	-----	-----	-----	-----	-----
19446	19781	11580	11448	2689	2666
-----	-----	-----	-----	-----	-----
2094	2106	1602	1589	3504	3735
-----	-----	-----	-----	-----	-----
3372	3414	2956	2966	1063	1122
-----	-----	-----	-----	-----	-----
1138	1111	2707	2764	5756	5729
-----	-----	-----	-----	-----	-----
1765	1764	27884	27982	6206	6239
-----	-----	-----	-----	-----	-----
4097	4085	390	<800	3403	3564
-----	-----	-----	-----	-----	-----
2949	2978	2569	2601	2216	2353
-----	-----	-----	-----	-----	-----
3174	3147	2228	2241	782	<800
-----	-----	-----	-----	-----	-----
1114	1111	3882	3909	4423	4633
-----	-----	-----	-----	-----	-----
247	2451	1295	1271	450	<800
-----	-----	-----	-----	-----	-----
15	<800	202	<800	601	<800
-----	-----	-----	-----	-----	-----
301	<800	7894	7858	3318	3413
-----	-----	-----	-----	-----	-----
6455	6537	504	<800	8046	8082
-----	-----	-----	-----	-----	-----
118	<800	782	<800	1110	1122
-----	-----	-----	-----	-----	-----
4564	4499	10447	10305	4632	4633
-----	-----	-----	-----	-----	-----
3599	3545	533	<800	477	<800
-----	-----	-----	-----	-----	-----
9664	9641	7717	7858	157	<800
-----	-----	-----	-----	-----	-----
250	<800	678	<800	5035	5074
-----	-----	-----	-----	-----	-----
15160	14802	1701	1643	761	<800
-----	-----	-----	-----	-----	-----
2025	1971	390	<800	441	<800
-----	-----	-----	-----	-----	-----
1905	1898	1946	1902	5627	5729
-----	-----	-----	-----	-----	-----
1412	1386	101	<800	7715	7683
-----	-----	-----	-----	-----	-----
3186	3147	637	<800	11604	11214
-----	-----	-----	-----	-----	-----
6226	6297	312	<800	2882	2697
-----	-----	-----	-----	-----	-----
4066	4085	2950	2966	3227	3413
-----	-----	-----	-----	-----	-----
1785	1764	1884	1902	4021	4204
-----	-----	-----	-----	-----	-----
133	<800	4779	4731	2721	2866
-----	-----	-----	-----	-----	-----
5971	5982	5304	5321	6232	6239
-----	-----	-----	-----	-----	-----
3070	3147	12363	12136	4125	4204

```

-----
620 <800 282 <800 2815 3011
-----
5882 5982 1170 1151 4140 4204
-----
4146 4085 4004 4067 8173 8082
-----
805 825 2339 2360 552 <800
-----
12533 12212 5102 5148 3690 3735
-----
5720 5773 21981 21990 14118 14343
-----
5473 5505 10346 10305 3444 3564
-----
117 <800 215 <800 6120 6239
-----
639 <800 784 813
-----
23 <800
-----
7100 7064
-----
FEATURES
source Location/Qualifiers
1..164500
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-229A12"
/clone_lib="RP11 human BAC library 11"
137635..137636
/note="1810 and 9-bp target site duplication excised from
between bases 137635 and 138973
CTGATGAATCCCTTAATGATTTGGTAAATAATCAATTAAGTGAATACATCT
TGTCAATGATCAATGTTGCGCAAAATCAATATCAGACACAAAGATGCGAA
Query Match 34.7% Score 43; DB 9; Length 164500;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 AGATGTCAGCTGTAACAAACAAGAGACTGTGTGTCT 89
Db 110158 AGATGTCAGCTGTAACAAACAAGAGACTGTGTGTCT 110116
RESULT 35
BD096974 706 bp DNA linear PAT 27-AUG-2002
BD096974
LOCUS SAG:apoptosis sensitivity gene.
DEFINITION SAG:apoptosis sensitivity gene.
ACCESSION BD096974
VERSION BD096974.1 GI:22642562
KEYWORDS JP 2001526063-A/9.
SOURCE JP 2001526063-A/9.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 706)
AUTHORS Sun, Y.
TITLE SAG:apoptosis sensitivity gene
JOURNAL Patent: JP 2001526063-A 9 18-DEC-2001;
WARNER LAMBERT CO
OS Unidentified
PN JP 2001526063-A/9
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PR 19-DEC-1997 US 60/068179, 11-SEP-1998 US 60/099840 PI
YI SUN
PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P7/02, A61P5/00,
A61P9/06,
PC A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C1201/68,
PC G01N33/50, G01N33/68, C12N15/00, A61K37/02, C12N5/00 CC
Strandedness: Double;

```

```

CC      Topology: Linear;
CC      SAG:apoptosis sensitivity gene
FH      Key      Location/Qualifiers
FT      CDS      1..291
FT      mat_peptide 1..291.
FEATURES
    source
        1..706
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
BASE COUNT      189 a      147 c      189 g      181 t
ORIGIN
Query Match      33.9%; Score 42; DB 6; Length 706;
Best Local Similarity 100.0%; Pred. No. 1.4e-14;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      83 GTGCTGTGGGAGATGTAATCATTCCTCCACAACTGCTGC 124
        |||||||
DB      175 GTGCTGTGGGAGATGTAATCATTCCTCCACAACTGCTGC 216

RESULT 36
AC0095698/c
LOCUS      AC0095698      254832 bp      DNA      linear      HTG_09-NOV-2002
DEFINITION      Rattus norvegicus clone CH230-9D13, WORKING DRAFT SEQUENCE.
ACCESSION      AC0095698
VERSION      AC0095698.6 GI:24817898
KEYWORDS      HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 254832)
REFERENCE
    Muzny, D., Marie, J., Metzker, M., Lee, J., Abramson, S., Adams, C., Alder, J.,
    Allen, C., Allen, H., Alsebrook, S., Amin, A., Anguitano, D.,
    Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
    Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
    Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
    Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
    Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
    Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
    Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
    Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dedertch, D.,
    Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
    Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
    Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
    Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
    Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
    Gebreyes, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
    Guneratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
    Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
    Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
    Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,
    Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
    Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
    Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
    Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
    Lorusshew, L., Louised, H., Lozado, R. J., Lu, X., Ma, J.,
    Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
    Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
    Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
    Milsavljeyevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
    Morgan, M., Morris, K., Morris, S., Muniasa, M., Murphy, M., Nat, L.,
    Nankervy, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
    Nwackeleneh, O., Okunonu, G., Olarnunsgoon, A., Pal, S., Parks, K.,
    Pasernak, S., Paul, H., Perez, A., Perez, A., Petrus, P., Frankoch, C.,
    Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L., L.,
    Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
    Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
    Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
    Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,

```

```

TITLE      JOURNAL
REFERENCE
    AUTHOR
    TITLE
    JOURNAL
REFERENCE
    AUTHOR
    TITLE
    JOURNAL
COMMENT
    Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
    of Molecular and Human Genetics, Baylor College of Medicine, One
    Baylor Plaza, Houston, TX 77030, USA
    3 (bases 1 to 254832)
    Rat Genome Sequencing Consortium.
    Direct Submission
    Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
    of Molecular and Human Genetics, Baylor College of Medicine, One
    Baylor Plaza, Houston, TX 77030, USA
    On Nov 9, 2002 this sequence version replaced gi:23267899.
    The sequence in this assembly is a combination of BAC based reads
    and whole genome shotgun sequencing reads assembled using Atlas
    (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
    in the feature table below represents a scaffold in the Atlas
    assembly (a 'contig-scaffold'). Within each contig-scaffold,
    individual sequence contigs are ordered and oriented, and separated
    by sized gaps filled with Ns to the estimated size. The sequence
    may extend beyond the ends of the clone and there may be sequence
    contigs within a contig-scaffold that consist entirely of whole
    genome shotgun sequence reads. Both ends and sequences and whole genome
    shotgun sequence only contigs will be indicated in the feature
    table.
    ----- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
    ----- Project Information
    Center project name: GDBC
    Center clone name: CH230-9D13
    ----- Summary Statistics
    Assembly program: Phrap; version 0.990329
    Consensus quality: 238918 bases at least Q40
    Consensus quality: 241069 bases at least Q30
    Consensus quality: 242304 bases at least Q20
    Estimated insert size: 243228; sum-of-contigs estimation
    Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
    -----
    * NOTE: Estimated insert size may differ from sequence length
    * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
    * NOTE: This is a 'working draft' sequence. It currently
    * consists of 1 contigs. Gaps between the contigs
    * are represented as runs of N. The order of the pieces
    * is believed to be correct as given, however the sizes
    * of the gaps between them are based on estimates that have
    * provided by the submitter.
    * This sequence will be replaced
    * by the finished sequence as soon as it is available and
    * the accession number will be preserved.
    *
    1 254832: contig of 254832 bp in length.
    Location/Qualifiers
    1..254832
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-9D13"
    1..1482
    misc_feature

```

```
misc_feature /note="wgs_end_extension
clone_end:Sp6"
2092..2921
/note="clone_boundary
clone_end:Sp6
site:EcORI
end_sequence:BH339916"
misc_feature /note="wgs_contig"
23357..235284
253718..254832
/note="wgs_contig"
complement(253801..254228)
/note="clone_boundary
clone_end:T7
site:EcORI
end_sequence:BH339912"
BASE COUNT 63080 a 58350 c 57410 g 64047 t 11945 others
ORIGIN

Query Match 29.8%; Score 37; DB 2; Length 254832;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 CCATCTGCAGGTCACAGTGATGATGCTGTCTTAG 48
Db 146354 CCATCTGCAGGTCACAGTGATGATGCTGTCTTAG 146318

RESULT 37
AL929565 120515 bp DNA linear ROD 16-APR-2003
LOCUS Mouse DNA sequence from clone RP23-454N16 on chromosome 4, complete
DEFINITION sequence.
ACCESSION AL929565.16 GI:30014183
VERSION AL929565.16
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 120515)
AUTHORS Andrew, R.
TITLE Direct SubMISSION
JOURNAL Submitted (16-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk
On Apr 16, 2003 this sequence version replaced gi:29825605.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
-----Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
```

```
Em, ENBL, Sw, SWISSPROT, Tr, TREMBL, Wp, WORMPEP, Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-454N16 is
from the RPI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.
FEATURES
source 1..120515
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-454N16"
/clone_lib="RPI-23"
BASE COUNT 34369 a 25840 c 26441 g 33865 t
ORIGIN

Query Match 29.0%; Score 36; DB 10; Length 120515;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 GCCATCTGCAGGTCACAGTGATGATGCTGTCTT 46
Db 57860 GCCATCTGCAGGTCACAGTGATGATGCTGTCTT 57825

RESULT 38
HSA323208 439 bp DNA linear PRI 18-JUL-2002
LOCUS Homo sapiens genomic sequence surrounding Not1 site, clone
DEFINITION NBI-073R.
ACCESSION AJ323208.1 GI:15867587
VERSION AJ323208.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 439)
AUTHORS Kutsenko, A.S., Gizaullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M.,
Podowski, R.M., Matubshkin, Y.G., Ganchandani, A., Muravenko, O.V.,
Levitsky, V.G., Kolchanov, N.A., Protchenov, A.I., Kashuba, V.I.,
Kiselev, L.L., Wasserman, W., Wahlestedt, C. and Zabarovsky, E.R.
Not1 flanking sequences: a tool for gene discovery and verification
of the human genome
Nucleic Acids Res. 30 (14), 3163-3170 (2002)
JOURNAL 22131767
MEDLINE 12136098
PUBMED 12136098
REFERENCE 2 (bases 1 to 439)
AUTHORS Zabarovsky, E.R.
TITLE Direct SubMISSION
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumourbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES
source 1..439
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NBI-073R"
BASE COUNT 63 a 161 c 146 g 68 t 1 others
ORIGIN

Query Match 28.2%; Score 35; DB 9; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATACGTCGCCCATCTGCAGGTCACAGTATG 35
Db 227 CCATACGTCGCCCATCTGCAGGTCACAGTATG 193
```

RESULT 39
AC137264/C
LOCUS
DEFINITION
AC137264 264606 bp DNA linear HTG 19-NOV-2002
Rattus norvegicus clone CH230-66B6, *** SEQUENCING IN PROGRESS ***,
3 unordered pieces.
AC137264
AC137264.1 GI:25073249
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 264606)
Rattus.
Muzny, D., Marie, Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Altschrocks, S., Amin, A., Angiano, D.,
Anyalelechi, V., Ayogi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Burch, P., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dedetich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dim, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagge, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Gervais, W.,
Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,
Gunnaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, D.,
Hernandez, R., Hines, S., Hilyk, S., Hume, J., Idlebird, D., Jackson, A.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Johnson, B., Johnson, R., Jollivet, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,
Karpeth, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyar, C.,
Kows, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshuler, L., Louie, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milošević, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nait, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackelme, O., Okunolu, G., Olarnpungsoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polindexter, A., Popovic, D., Pritus, E., Pu, L.,
Rizzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.U.,
Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sheed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabori, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Woodson, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstein, G., and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 264606)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KCMN
Center clone name: CH230-66B6
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 21074/ bases at least Q40
Consensus quality: 21494/ bases at least Q30
Consensus quality: 218032/ bases at least Q20
Estimated insert size: 219724/ sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Gendank_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 66491: contig of 66491 bp in length
* 66492 66591: gap of unknown length
* 66592 109095: contig of 42504 bp in length
* 109096 109195: gap of unknown length
* 109196 264606: contig of 15411 bp in length.
* Location/Qualifiers
* 1. 264606
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-66B6"
* 66592 . 67666
* /note="wgs contig"
* 82489 . 84829
* /note="wgs contig"
BASE COUNT 5611 a 53481 c 54146 g 56734 t 44134 others
ORIGIN
Query Match 28.2%; Score 35; DB 2; Length 264606;
Best Local Similarity 100.0%; Pred. No. 4.9e-10;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 GGATCCCTGCTTACATGTCAGCTGAAAACAAAC 68
DB 158502 GGATCCCTGCTTACATGTCAGCTGAAAACAAAC 158468
RESULT 40
BC011127
LOCUS BC011127 1125 bp mRNA linear ROD 16-APR-2003
DEFINITION Mus musculus ring finger protein 7, mRNA (cDNA clone MGC:19076
IMAGE:4194107), complete cds.
ACCESSION BC011127
VERSION BC011127.1 GI:15029807
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1125)

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D., Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stajich T.E., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J., Bosak S.A., McEwan P.J., Mckernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Boultard G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalske U., Smalls D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
MEDLINE
PUBMED
22388257
12477932

REFERENCE
TITLES
2 (bases 1 to 1125)

Strausberg R.
Direct Submission
Submitted (25-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne P.H., Garcia A.M., Lu X., Hulyk S.W., Louesged H., Kowis C.R., Sneed A.J., Martin R.G., Muzny D.M., Nanavati A.N., Gibbs R.A.

FEATURES
source
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 24 Row: k Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, similarity but not identity to protein.
Location/Qualifiers
1..1125
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:19076 IMAGE:4194107"
/tissue_type="Liver, normal. 5 month old male mouse."
/clone_1lb="NCI CGAP_L19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1125
/gene="Rnf7"
/note="synonym: SAG"
/db_xref="locusid:19823"
/db_xref="MGI:1337096"
12..353
/codon_start=1
/product="Rnf7 protein"

gene
CDS

Query Match
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
11 GCCATCTGCAGGCTCAGGTGATGATGCTG 42
162 GCCATCTGCAGGCTCAGGTGATGATGCTG 193

RESULT 41
BD096966 1140 bp DNA linear PAT 27-AUG-2002
DEFINITION
SAG:apoptosis sensitivity gene.
ACCESSION
BD096966
VERSION
BD096966.1 GI:22642554
KEYWORDS
JP 2001526063-A/1.
SOURCE
unidentified
unclassified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1140)
Sun, Y.
SAG:apoptosis sensitivity gene
Patent: JP 2001526063-A 1 18-DEC-2001;
WARNER LAMBERT CO
OS Unidentified
PN JP 2001526063-A/1
PD 18-DEC-2001
PF 15-DEC-1998 JP 2000525451
PI 19-DEC-1997 US 60/068179, 11-SEP-1998 US 60/099840
YI SUN
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P17/02,A61P35/00,
PC A61P39/06,
PC A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC
C12N5/10,C1201/68,
PC G01N33/50,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC /note = 'Mouse SAG';
FH Key Location/Qualifiers
FT CDS 17..355
FT mat_peptide 17..355
FT misc_feature 1..1140.
Location/Qualifiers
1..1140
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT
ORIGIN
302 a 224 c 287 g 327 t

Query Match
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
11 GCCATCTGCAGGCTCAGGTGATGATGCTG 42
167 GCCATCTGCAGGCTCAGGTGATGATGCTG 198

RESULT 42
AF092877 1140 bp mRNA linear ROD 24-JUL-2001
LOCUS
AF092877
DEFINITION
Mus musculus zinc RING finger protein SAG mRNA, complete cds.
ACCESSION
AF092877

```

VERSION      AF092877.1 GI:4586031
KEYWORDS
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
REFERENCE    1 (bases 1 to 1140)
AUTHORS      Duan,H., Wang,Y., Aviram,M., Swaroop,M., Loo,J.A., Bian,J.,
              Tian,Y., Mueller,T., Bisgaier,C.L. and Sun,Y.
              SAG, a novel zinc RING finger protein that protects cells from
              apoptosis induced by redox agents
              Mol. Cell. Biol. 19 (4), 3145-3155 (1999)
TITLE        JOURNAL
              MEDLINE
              PUBMED
              1082581
              2 (bases 1 to 1140)
              Sun,Y.
              Alterations of SAG mRNA in human cancer cell lines: requirement for
              the RING finger domain for apoptosis protection
              Carcinogenesis 20 (10), 1899-1903 (1999)
              99435944
              10506102
              3 (bases 1 to 1140)
              Swaroop,M., Bian,J., Aviram,M., Duan,H., Bisgaier,C.L., Loo,J.A.
              and Sun,Y.
              Expression, purification, and biochemical characterization of SAG,
              a RING finger redox-sensitive protein
              Free Radical Biol. Med. 27, 193-202 (1999)
              4 (bases 1 to 1140)
              Swaroop,M., Wang,Y., Miller,P., Duan,H., Jatkoe,T., Madore,S.J. and
              Sun,Y.
              Yeast homolog of human SAG/ROC2/Rbx2/Hrt2 is essential for cell
              growth, but not for germination: chip profiling implicates its role
              in cell cycle regulation
              Oncogene 19 (24), 2855-2866 (2000)
              10851089
              5 (bases 1 to 1140)
              Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R.,
              Kung,H.F., Zhang,H. and Sun,Y.
              Promotion of S-phase entry and cell growth under serum starvation
              by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component:
              association with inhibition of p27 accumulation
              Mol. Cell. Biol. 20 (1), 37-46 (2001)
              21152847
              11255262
              6 (bases 1 to 1140)
              Sun,Y.
              Direct Submission
              Submitted (16-SEP-1998) Department of Molecular Biology,
              Parke-Davis, 2800 Plymouth Rd, Ann Arbor, MI 48105, USA
              1. .1140
              Location/Qualifiers
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6XCB"
              /db_xref="taxon:10090"
              /sex="female"
              /tissue_type="lung"
              /dev_stage="6-8 weeks"
              /note="redox sensitive, metal binding; expression protects
              cells from apoptosis induced by redox compounds"
              /codon_start=1
              /product="zinc RING finger protein SAG"
              /protein_id="AAD25961.1"
              /db_xref="GI:4586032"
              /translation="MADVEDGEPCVLSHSGSAGSKSGDMPFLTKKNVAVMSWD
              VECDFCAICRVQVMDACLRCAENKQECVAVWGECNHSFNCCMSLWVKNRCPLC
              QQDWVYRIK"
              302 a 224 c 287 g 327 t
              BASE COUNT
              ORIGIN
              Query Match
              25.8%; Score 32; DB 10; Length 1140;

```

```

Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 GCCATCTGCAGGATCCAGTATGATGATCCCTG 42
Db 167 GCCATCTGCAGGATCCAGTATGATGATCCCTG 198
RESULT 43
AC123935
LOCUS       AC123935
DEFINITION Mus musculus chromosome 13 clone RP23-354J3, complete sequence.
ACCESSION  AC123935
VERSION    AC123935.4 GI:23462972
KEYWORDS   HTG.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 219370)
AUTHORS    Mckherson,J.D. and Waterston,R.H.
TITLE      The sequence of Mus musculus clone
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 219370)
AUTHORS    Mckherson,J.D. and Waterston,R.H.
TITLE      Direct Submission
JOURNAL    Submitted (05-JUN-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
              3 (bases 1 to 219370)
              Mckherson,J.D. and Waterston,R.H.
              Direct Submission
              Submitted (13-SEP-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
              4 (bases 1 to 219370)
              Mckherson,J.D. and Waterston,R.H.
              Direct Submission
              Submitted (02-OCT-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
              On Oct 2, 2002 this sequence version replaced gi:22748547.
              COMMENT
              ----- Genome Center -----
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu/gsc/index.shtml
              Contact: submissions@watson.wustl.edu
              ----- Project Information -----
              Center project name: M BA0354J03
              -----
              FEATURES
              source
              1. .219370
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10090"
              /chromosome="13"
              /clone="RP23-354J3"
              BASE COUNT  61709 a 45568 c 45395 g 66698 t
              ORIGIN
              Query Match
              25.8%; Score 32; DB 10; Length 219370;
              Best Local Similarity 100.0%; Pred. No. 3.7e-08;
              Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
              QY 11 GCCATCTGCAGGATCCAGTATGATGATCCCTG 42
              Db 79023 GCCATCTGCAGGATCCAGTATGATGATCCCTG 79054
              RESULT 44
              BV071691
              LOCUS       BV071691
              DEFINITION S212P6546FB3.T0 CZECHII/E1 Mus musculus SFS genomic, sequence
              tagged site.
              ACCESSION BV071691
              612 bp DNA linear STS 31-MAY-2003

```



```

* 3512 3611: gap of 100 bp
* 3612 4746: contig of 1135 bp in length
* 4747 4846: gap of 100 bp
* 4847 6030: contig of 1184 bp in length
* 6031 6130: gap of 100 bp
* 6131 7290: contig of 1160 bp in length
* 7291 7390: gap of 100 bp
* 7391 8577: contig of 1187 bp in length
* 8578 8677: gap of 100 bp
* 8678 9937: contig of 1260 bp in length
* 9938 10037: gap of 100 bp
* 10038 11192: contig of 1155 bp in length
* 11193 11292: gap of 100 bp
* 11293 12454: contig of 1162 bp in length
* 12455 12554: gap of 100 bp
* 12555 13718: contig of 1164 bp in length
* 13719 13818: gap of 100 bp
* 13819 15002: contig of 1184 bp in length
* 15003 15102: gap of 100 bp
* 15103 16234: contig of 1132 bp in length
* 16235 16334: gap of 100 bp
* 16335 17461: contig of 1127 bp in length
* 17462 17561: gap of 100 bp
* 17562 18650: contig of 1089 bp in length
* 18651 18750: gap of 100 bp
* 18751 19912: contig of 1162 bp in length
* 19913 20013: gap of 100 bp
* 20013 21199: contig of 1187 bp in length
* 21200 21299: gap of 100 bp
* 21300 22462: contig of 1163 bp in length
* 22463 22562: gap of 100 bp
* 22563 23720: contig of 1158 bp in length
* 23721 23820: gap of 100 bp
* 23821 25005: contig of 1185 bp in length
* 25006 25105: gap of 100 bp
* 25106 26265: contig of 1160 bp in length
* 26266 26365: gap of 100 bp
* 26366 27528: contig of 1163 bp in length
* 27529 27628: gap of 100 bp
* 27629 28788: contig of 1160 bp in length
* 28789 28888: gap of 100 bp
* 28889 30034: contig of 1146 bp in length
* 30035 30134: gap of 100 bp
* 30135 31236: contig of 1102 bp in length
* 31237 31336: gap of 100 bp
* 31337 32512: contig of 1176 bp in length
* 32513 32612: gap of 100 bp
* 32613 33813: contig of 1201 bp in length
* 33814 33913: gap of 100 bp
* 33914 35107: contig of 1194 bp in length
* 35108 35207: gap of 100 bp
* 35208 36431: contig of 1224 bp in length
* 36432 36531: gap of 100 bp
* 36532 37701: contig of 1170 bp in length
* 37702 38967: contig of 1166 bp in length
* 38968 39067: gap of 100 bp
* 39068 40235: contig of 1168 bp in length
* 40236 40335: gap of 100 bp
* 40336 41492: contig of 1157 bp in length
* 41493 41592: gap of 100 bp
* 41593 42691: contig of 1099 bp in length
* 42692 42791: gap of 100 bp
* 42792 43970: contig of 1179 bp in length
* 43971 44070: gap of 100 bp
* 44071 45150: contig of 1080 bp in length
* 45151 45250: gap of 100 bp
* 45251 46347: contig of 1097 bp in length
* 46348 46447: gap of 100 bp
* 46449 47608: contig of 1161 bp in length
* 47609 47708: gap of 100 bp
* 47709 48881: contig of 1173 bp in length
* 48882 48981: gap of 100 bp

```

```

* 48982 50240: contig of 1259 bp in length
* 50241 50340: gap of 100 bp
* 50341 51544: contig of 1204 bp in length
* 51545 51644: gap of 100 bp
* 51645 52828: contig of 1184 bp in length
* 52829 52929: gap of 100 bp
* 52929 54157: contig of 1229 bp in length
* 54158 54257: gap of 100 bp
* 54258 55450: contig of 1193 bp in length
* 55451 55550: gap of 100 bp
* 55551 56731: contig of 1181 bp in length
* 56732 56831: gap of 100 bp
* 56832 58019: contig of 1188 bp in length
* 58020 58119: gap of 100 bp
* 58120 59285: contig of 1166 bp in length
* 59286 59385: gap of 100 bp
* 59386 60525: contig of 1140 bp in length
* 60526 60625: gap of 100 bp
* 60626 61717: contig of 1092 bp in length
* 61718 61817: gap of 100 bp
* 61818 62991: contig of 1174 bp in length
* 62992 63091: gap of 100 bp
* 63092 64267: contig of 1176 bp in length
* 64268 64367: gap of 100 bp
* 64368 65552: contig of 1185 bp in length
* 65553 65652: gap of 100 bp
* 65653 66853: contig of 1201 bp in length
* 66854 66953: gap of 100 bp
* 66954 68145: contig of 1192 bp in length
* 68146 68245: gap of 100 bp
* 68246 69423: contig of 1178 bp in length
* 69424 69523: gap of 100 bp
* 69524 70737: contig of 1214 bp in length
* 70738 70837: gap of 100 bp
* 70838 71997: contig of 1160 bp in length
* 71998 72097: gap of 100 bp
* 72098 72554: contig of 1157 bp in length

```

Query Match 25.0%; Score 31; DB 2; Length 112321;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 CTGCAGGCTCCAGGTGATGATGCTGCTT 46
 Db 70515 CTGCAGGCTCCAGGTGATGATGCTGCTT 70545

RESULT 46
 AC103666/c
 LOCUS Mus musculus clone RP23-397H13, LOW-PASS SEQUENCE SAMPLING.
 AC103666
 AC103666.2 GI:27311471
 DEFINITION HTG: HTGS PHASED.
 ACCESSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 112321)
 2 (bases 1 to 112321)
 REFERENCE
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgeater, B.,
 Brown, A., Camarata, J., Campoliano, A., Chang, J., Chararo, B.,
 Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Darello, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gargyala, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K.,
 Lamazares, R., Landers, T., Lehocsky, J., Levine, R., Liu, G.,

Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
 Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Reta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Struss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travs, N., Triglio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (29-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 113321)

Barra, N., Busbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Birtner, B., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gadyana, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamet, A., Karakas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,
 Mathews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T.,
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunhkhang, P., Pierre, N., Raymond, C., Reta, R.,
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (20-DEC-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Dec 20, 2002 this sequence version replaced gi:1719550.
 All repeats were identified using RepeatMasker:
 Smith, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L19615
 Center clone name: 397_H_13

* NOTE: This record contains 89 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1	1147: contig of 1147 bp in length	7391	8577: contig of 1187 bp in length
1148	1247: gap of 100 bp	8578	8677: gap of 100 bp
1248	1248: contig of 1101 bp in length	8678	9937: contig of 1260 bp in length
2349	2448: gap of 100 bp	9938	10037: gap of 100 bp
2449	3511: contig of 1063 bp in length	10038	11192: contig of 1155 bp in length
3512	3611: gap of 100 bp	11193	11292: gap of 100 bp
3612	4746: contig of 1135 bp in length	11293	12454: contig of 1162 bp in length
4747	4846: gap of 100 bp	12455	12554: gap of 100 bp
4847	6030: contig of 1184 bp in length	12555	13718: contig of 1164 bp in length
6031	6130: gap of 100 bp	13719	13818: gap of 100 bp
6131	7290: contig of 1160 bp in length	13819	15002: contig of 1184 bp in length
7291	7390: gap of 100 bp	15003	15102: gap of 100 bp
		15103	16234: contig of 1132 bp in length
		16235	16334: gap of 100 bp
		16335	17461: contig of 1137 bp in length
		17462	17561: gap of 100 bp
		17562	18650: contig of 1089 bp in length
		18651	18750: gap of 100 bp
		18751	19912: contig of 1162 bp in length
		19913	20012: gap of 100 bp
		20013	21199: contig of 1187 bp in length
		21200	21299: gap of 100 bp
		22462	22462: contig of 1163 bp in length
		22463	23562: gap of 100 bp
		23563	23720: contig of 1158 bp in length
		23721	23820: gap of 100 bp
		23821	25005: contig of 1185 bp in length
		25006	25105: gap of 100 bp
		25106	26265: contig of 1160 bp in length
		26266	26365: gap of 100 bp
		26366	27528: contig of 1163 bp in length
		27529	27628: gap of 100 bp
		27629	28788: contig of 1160 bp in length
		28789	28888: gap of 100 bp
		28889	30034: contig of 1146 bp in length
		30035	30134: gap of 100 bp
		30135	31236: contig of 1102 bp in length
		31237	31336: gap of 100 bp
		31337	32512: contig of 1176 bp in length
		32513	32612: gap of 100 bp
		32613	33813: contig of 1201 bp in length
		33814	33913: gap of 100 bp
		33914	35107: contig of 1194 bp in length
		35108	35207: gap of 100 bp
		35208	36431: contig of 1224 bp in length
		36432	36531: gap of 100 bp
		36532	37701: contig of 1170 bp in length
		37702	37801: gap of 100 bp
		37802	38967: contig of 1166 bp in length
		38968	39067: gap of 100 bp
		39068	40235: contig of 1168 bp in length
		40236	40335: gap of 100 bp
		40336	41492: contig of 1157 bp in length
		41493	41592: gap of 100 bp
		41593	42691: contig of 1099 bp in length
		42692	42792: gap of 100 bp
		42792	43970: contig of 1179 bp in length
		43971	44070: gap of 100 bp
		44071	45150: contig of 1080 bp in length
		45151	45250: gap of 100 bp
		45251	46347: contig of 1097 bp in length
		46348	46447: gap of 100 bp
		46448	47608: contig of 1161 bp in length
		47609	47708: gap of 100 bp
		47709	48881: contig of 1173 bp in length
		48882	48981: gap of 100 bp
		48982	50240: contig of 1259 bp in length
		50241	50340: gap of 100 bp
		50341	51544: contig of 1204 bp in length
		51545	51644: gap of 100 bp
		51645	52828: contig of 1184 bp in length
		52829	52928: gap of 100 bp
		52929	54157: contig of 1229 bp in length


```

misc_feature      3313..6908
                  /note="assembly_fragment"
misc_feature      7009..14434
                  /note="assembly_fragment"
misc_feature      14535..32181
                  /note="assembly_fragment"
misc_feature      32282..51167
                  /note="assembly_fragment"
misc_feature      51268..72870
                  /note="assembly_fragment"
misc_feature      72971..149992
                  /note="assembly_fragment"
misc_feature      150093..187353
                  /note="assembly_fragment"
                  /note="assembly_fragment"
                  clone_end:T7
vector_side:right"
BASE COUNT      60628 a 34855 c 36129 g 54836 t 905 others
ORIGIN

```

```

Query Match      25.0%; Score 31; DB 2; Length 187353;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      11 GCCATCTGCAGGCTCCAGGTGATGATGCCT 41
         |||||
Db      46893 GCCATCTGCAGGCTCCAGGTGATGATGCCT 46923

```

```

RESULT 48      148290 bp      DNA      linear      HTG 11-JUN-2002
AC119145      Rattus norvegicus chromosome 1 clone RP32-4G14, WORKING DRAFT
LOCUS
DEFINITION
SEQUENCE, 10 unordered pieces.
AC119145
AC119145.1 GI:20303534
VERSION
HTG; HTGS_PHASE1; HTGS_FULFILLTOP; HTGS_DRAFT.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 148290)
Taenzer,S., Monti,J., Gloeckner,G., Goesele,C., Baumgart,C.,
Huebner,N. and Platzzer,M.
Rat chromosome 1 genomic sequence
Unpublished
2 (bases 1 to 148290)
Taenzer,S. and Platzzer,M.
Direct Submision
Submitted (25-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
3 (bases 1 to 148290)
Taenzer,S. and Platzzer,M.
Direct Submision
Submitted (11-JUN-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany

```

```

COMMENT
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: r308
Center clone name: RP32-4G14
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990139
Consensus quality: 140730 bases at least Q40
Consensus quality: 143676 bases at least Q30
Consensus quality: 145595 bases at least Q20
Quality coverage: 8.49
-----

```

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

***** NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1      25364: contig of 25364 bp in length
*      25365      25464: gap of unknown length
*      25465      38093: contig of 12629 bp in length
*      38094      38193: gap of unknown length
*      38194      83797: contig of 45604 bp in length
*      83798      83897: gap of unknown length
*      83898      90673: contig of 6776 bp in length
*      90674      90773: gap of unknown length
*      90774      107588: contig of 16815 bp in length
*      107589      107688: gap of unknown length
*      107689      112170: contig of 4482 bp in length
*      112171      112770: gap of unknown length
*      112771      119805: contig of 7535 bp in length
*      119806      119905: gap of unknown length
*      119906      137949: contig of 18044 bp in length
*      137950      138049: gap of unknown length
*      138050      146503: contig of 8454 bp in length
*      146504      146603: gap of unknown length
*      146604      148290: contig of 1667 bp in length.
Location/Qualifiers
1..148290
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/chromosome="1"
/clone="RP32-4G14"
/clone_lib="RPCT-32 BAC Library"

```

```

BASE COUNT      41779 a 31830 c 31870 g 41911 t 900 others
ORIGIN

```

```

Query Match      24.2%; Score 30; DB 2; Length 148290;
Best Local Similarity 100.0%; Pred. No. 6.4e-07;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      8 TCCGCCATCTGCAGGCTCCAGGTGATGCAT 37
         |||||
Db      17405 TCCGCCATCTGCAGGCTCCAGGTGATGCAT 17434

```

```

RESULT 49      189541 bp      DNA      linear      HTG 20-NOV-2002
AC118520/c      Rattus norvegicus clone CH230-30421, *** SEQUENCING IN PROGRESS
LOCUS
DEFINITION
*** 2 unordered pieces.
AC118520
AC118520.4 GI:25138134
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 189541)
Muzny,D.,Marie., Metzker,M.,Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angilano,D.,

```

Ayvalidebechi, V., Aoyagi, A., Ayvodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Bunay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dedrich, D.,
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Dublin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabits, A., Gante, R., Garcia, A., Garner, T., Gava, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guvavara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladin, S. L., Hodgson, A., Hognes, M.,
 Hollins, B., Howells, S., Huliy, S., Hume, J., Idelberger, D., Jackson, A.,
 Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, C.,
 Kwis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, Y., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshewa, L., Lounsged, H., Loxado, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmud, M., Malloy, K., Mangum, A.,
 Mangun, B., Magna, P., Martin, K., Martin, R., Mattheisz, E.,
 Mahoney, S., McLeod, M. P., McNeill, T. Z., Meene, E.,
 Mlosoaviljevic, A., Miner, G., Mina, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
 Narkervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nockemleh, O., Okunou, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poldexier, A., Popovic, D., Primus, E., Pu, L., L.,
 Puzio, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Rella, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X. Z., Soreller, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Swatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmail, K.,
 Valas, R., Varga, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willison, R., Wleciyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 189541)
 Direct Submission
 Submitted (18-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 189541)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 20, 2002 this sequence version replaced gi:123914484.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature

table.	----- Genome Center
	Center: Baylor College of Medicine
	Center code: BCM
	Web site: http://www.hgsc.bcm.tmc.edu/
	Contact: hgsc-help@bcm.tmc.edu
	----- Project Information
	Center project name: GUNO
	Center clone name: CH230-304J21
	----- Summary Statistics
	Assembly program: Phrap; version 0.990329
	Consensus quality: 173927 bases at least Q40
	Consensus quality: 175964 bases at least Q30
	Consensus quality: 176952 bases at least Q20
	Estimated insert size: 178045; sum-of-coverage estimation
	Quality coverage: 6x in Q20 bases; sum-of-coverage estimation

* NOTE:	Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).	
* NOTE:	This is a "working draft" sequence. It currently
* consists of 2 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
* 1 188414: contig of 188414 bp in length	
* 188415 188514: gap of unknown length	
* 188515 189541: contig of 1027 bp in length.	
Location/Qualifiers	
1. 189541	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10116"	
/clone="CH230-304J21"	
121774. 123263	
/note="wgs contig"	
/note="complement[184170.. 185020]"	
/note="Clone_boundary	
clone_end:17	
site:	
end_sequence:B2225638"	
187321. 188414	
/note="wgs end_extension	
clone_end:17"	
BASE COUNT 49137 a 39834 c 38226 g 50535 t 11809 others	
ORIGIN	
Query Match 24.2%; Score 30; DB 2; Length 189541;	
Best Local Similarity 100.0%; Pred. No. 6,5e-07;	
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
GC 73 GGACTGTTGTGCTCTGGGAGATGTAA 102	
Db 93758 GGACTGTTGTGCTCTGGGAGATGTAA 93729	
RESULT 50	
AC135657 198756 bp DNA linear HTG 19-NOV-2002	
LOCUS	
DEFINITION Rattus norvegicus clone CH230-416B17, WORKING DRAFT SEQUENCE.	
AC135657	
AC135657.2 GI:25073392	
VERSION	
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULFILLTOP.	
SOURCE	
ORGANISM Rattus norvegicus (Norway rat)	
Rattus norvegicus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
Rattus.	
1 (bases 1 to 198756)	
Murphy,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,	
Allen,C., Allen,H., Albrooks,S., Amth,A., Angiano,D.,	

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhmed, F., Bissolo, K., Blair, J., Blankenburg, K., Blych, P., Brown, M., Bryant, K., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceesar, H., Center, A., Chacko, U., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hu, Y., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Loulseged, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Mijiga, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelam, O., Okunolu, G., Olarunpogun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shattman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soes, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usman, K., Vals, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Woodson, H., Wortley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Weinstock, G., and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 198756)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 198756)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:24158510.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBVZ
Center clone name: CH230-416B17
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 187507 bases at least Q40
Consensus quality: 189483 bases at least Q30
Consensus quality: 190529 bases at least Q20
Estimated insert size: 192210; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 198756: contig of 198756 bp in length.
Location/Qualifiers
1..198756
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-416B17"
1..1088
/note="wgs end extension
clone_end:5p6"
misc_feature
1849..3549
/note="wgs end-extension
clone_end:5p6"
5730..6605
/note="clone boundary
clone_end:5p6
site:
end_sequence:B2197454"
complement(191645..192594)
/note="clone boundary
clone_end:T7"
misc_feature
end_sequence:B2197451"
197550..198756
/note="wgs end-extension
clone_end:T7"
BASE COUNT 51425 a 44460 c 43469 g 51998 t 7404 others
ORIGIN
Query Match 24.2%; Score 30; DB 2; Length 198756;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 AGATGTCAAGCTGAAACAAACAAAGAGAC 76
DB 176453 AGATGTCAAGCTGAAACAAACAAAGAGAC 176482
Search completed: November 7, 2003, 13:29:07
Job time : 825.044 secs

```

THIS PAGE BLANK (USPTO)